

Harvis, A.
10/726160
Seq. IDs 1 & 2

GenCore version 5.1.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2006, 05:53:08 ; Search time 15911 Seconds
(without alignments)
11377.987 Million cell updates/sec

Title: US-10-726-160-1
Perfect score: 2831
Sequence: 1 ggacgagcggagtggag.....atgaaaaaaaaaaaaaa 2831

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 6366136 seqs, 31973710525 residues

Total number of hits satisfying chosen parameters: 12732272

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_env:*
2: gb_pat:*
3: gb_ph:*
4: gb_pl:*
5: gb_pr:*
6: gb_ro:*
7: gb_sts:*
8: gb_sy:*
9: gb_un:*
10: gb_vi:*
11: gb_ov:*
12: gb_htg:*
13: gb_in:*
14: gb_om:*
15: gb_ba:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2831	100.0	2831	5	AF345896	AF345896 Homo sapi
2	2816.2	99.5	2838	2	AX213281	AX213281 Sequence
3	2804	99.0	4422	2	CQ413115	CQ413115 Sequence
4	2804	99.0	4422	2	CQ493832	CQ493832 Sequence
5	2804	99.0	4422	2	AX188263	AX188263 Sequence
6	2802.4	99.0	3008	5	AK000742	AK000742 Homo sapi
7	2802.2	99.0	2969	5	BC033540	BC033540 Homo sapi
8	2802	99.0	4450	2	AX780410	AX780410 Sequence
9	2789.8	98.5	4221	2	AX926533	AX926533 Sequence
10	2789.8	98.5	4221	2	AX951786	AX951786 Sequence
11	2789.8	98.5	4221	2	AX960064	AX960064 Sequence
12	2789.8	98.5	4221	2	CS033737	CS033737 Sequence
13	2789.8	98.5	4221	2	CS042689	CS042689 Sequence
14	2789.8	98.5	4221	5	AF195765	AF195765 Homo sapi
15	2775.6	98.0	2782	5	BC033297	BC033297 Homo sapi
16	2290.6	80.9	3685	2	BD159852	BD159852 Primer fo
17	2290.6	80.9	3685	2	AX882660	AX882660 Sequence
18	2290.6	80.9	3685	5	AK027651	AK027651 Homo sapi

19	1833	64.7	1833	2	BD156087	BD156087 Primer fo
20	1833	64.7	1833	2	AX876358	AX876358 Sequence
21	1833	64.7	1833	5	AK001261	AK001261 Homo sapi
22	1812.6	64.0	2442	6	AB095735	AB095735 Mus muscu
23	1783.8	63.0	4160	6	BC060208	BC060208 Mus muscu
24	1729.2	61.1	3175	6	AB095736	AB095736 Mus muscu
25	1520	53.7	1871	2	AX274866	AX274866 Sequence
26	1490.4	52.6	1803	2	CQ728118	CQ728118 Sequence
27	1273.8	45.0	1291	2	BD156017	BD156017 Primer fo
28	1273.8	45.0	1291	2	AX876224	AX876224 Sequence
29	1273.8	45.0	1291	5	AK001206	AK001206 Homo sapi
30	856.4	30.3	905	2	BD147262	BD147262 Primer fo
31	856.4	30.3	905	2	AX867200	AX867200 Sequence
32	843	29.8	3017	11	AJ720316	AJ720316 Gallus ga
33	834.2	29.5	160200	5	AL606468	AL606468 Human DNA
34	804.4	28.4	853	2	BD147403	BD147403 Primer fo
35	804.4	28.4	853	2	AX867341	AX867341 Sequence
36	680.6	24.0	215416	12	AC153228	AC153228 Bos tauru
37	680.6	24.0	232073	12	AC171797	AC171797 Bos tauru
38	652.2	23.0	2904	11	BC097560	BC097560 Xenopus l
39	649.8	23.0	205300	6	AC132830	AC132830 Mus muscu
40	649.4	22.9	2639	11	CR848464	CR848464 Xenopus t
41	649.4	22.9	2676	11	BC064851	BC064851 Xenopus t
42	633.6	22.4	2778	11	BC073015	BC073015 Xenopus l
43	625.4	22.1	242922	12	AC105627	AC105627 Rattus no
44	581.6	20.5	1309	11	CR386886	CR386886 Gallus ga
45	558	19.7	708	2	BD129416	BD129416 Human gen

ALIGNMENTS

RESULT 1	AF345896	Homo sapiens RA-regulated nuclear matrix-associated protein mRNA, complete cds.	2831 bp	mRNA	linear	PRI 21-MAY-2001
AF345896	LOCUS	AF345896	2831 bp	mRNA	linear	PRI 21-MAY-2001
DEFINITION	AF345896.1	GI:14161319				
ACCESSION	AF345896					
VERSION	AF345896.1	GI:14161319				
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
PUBMED						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
FEATURES						
source						
CDS						

ETGVPPPPFCTSSAPNHEHVLAVANEBCFVRLNLTNSQSFRRKKPEWMAHWNVA
DLAWPGLKLTADGQTAQKVDKAGLIGTICGHQCSKLKSVAFSKLAPVCTGQ
RDNGNIVKCNKGQFQVQVNLISGAHNTSDGTQSPKKPKKSKFAPVFOQS
VTVVLFQDENTLVASAGAVGIIKWMDRLKLVIAEQBSKSLVPGSSTKGLGYS
LILDSFQSTLACTMDNTYMNFTLKTSPVAIEHQNSHTVFKVSLSPDQFLVS
GSDSEAIWKVSTPQWPLTGLHSHQSVTWCSPDGTQKATCSDDNTLTKLWLRN
GLEEKCDGKLTVSGVASOKKVESPLVTVTSOSTAKAPVKKVCNNSPSSAAC
APSCAGDLPLPSNTPFISITSPAKARSPIINRGRSVSVSPKPSFKMGRNWT
PSSSPPIIPPASETGVNMLPRKALIPVQSKOCSAGSSRNVRKRLDSSCLSEVKOK
FASCNVCTBEDGVNHLICLJAGNQDSKQSLKSGTSTKIGAGTISIEPPSP
ISPVNSRSCGTLPLRPLCGEBSVMVKGKNSLPHKNSMLAAAKRAKSNPSPSPS
OTPNRROSCGTLPLVPVITPSSMKICTVPHKSOEDFCCEPHSTNIA"

ORIGIN

Query Match	100.0%;	Score 2831;	DB 5;	Length 2831;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2831;	Conservative	0;	Mismatches	0;
			Indels	Gaps
Qy	1	GGCAGACGGGAGTTGGAGGCGATAACGATTTGTTGTGTGAGAGCGCAACGTCGCGATT	60	
Db	1	GGCAGACGGGAGTTGGAGGCGATAACGATTTGTTGTGTGAGAGCGCAACGTCGCGATT	60	
Qy	61	TCTGCTGAACCTTGGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACC	120	
Db	61	TCTGCTGAACCTTGGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACC	120	
Qy	121	CTGATGCTCTTCAATTCGGTGCTCCGCAGCCCCAGCTTGGCGTCTCGAGAAATCGATGG	180	
Db	121	CTGATGCTCTTCAATTCGGTGCTCCGCAGCCCCAGCTTGGCGTCTCGAGAAATCGATGG	180	
Qy	181	TCTTCACAATACCCCTCTTCAATCCCTTCTGACTGGTTATCAGTCAGTGGTAAATGATGAA	240	
Db	181	TCTTCACAATACCCCTCTTCAATCCCTTCTGACTGGTTATCAGTCAGTGGTAAATGATGAA	240	
Qy	241	CACACTTCTTATGGAGAAACAGGAGTCCAGTTCCCTCTCTTTGGATGTACCTTCTCTCT	300	
Db	241	CACACTTCTTATGGAGAAACAGGAGTCCAGTTCCCTCTCTTTGGATGTACCTTCTCTCT	300	
Qy	301	GCTCCCAATATGGAACATGTACTAGCAGTTGGCAATGAAGAAGGCTTTGTTCCGATTTGAT	360	
Db	301	GCTCCCAATATGGAACATGTACTAGCAGTTGGCAATGAAGAAGGCTTTGTTCCGATTTGAT	360	
Qy	361	AACACAGAAATCACAAAGTTTTCAGAAAGAGTGCTTCAAGAATGGATGGCTCACTGGAAAT	420	
Db	361	AACACAGAAATCACAAAGTTTTCAGAAAGAGTGCTTCAAGAATGGATGGCTCACTGGAAAT	420	
Qy	421	GCGGCTTTTGACCTCGGCTGGGTTCCCTGGTGAACTTAAACTTTGTACAGCAGCAGTGAT	480	
Db	421	GCGGCTTTTGACCTCGGCTGGGTTCCCTGGTGAACTTAAACTTTGTACAGCAGCAGTGAT	480	
Qy	481	CAAAACAGCCAAATTTTGGGACGTAAAGCTGGTGAGCTGATTTGGAACATGCAAAAGGTCAT	540	
Db	481	CAAAACAGCCAAATTTTGGGACGTAAAGCTGGTGAGCTGATTTGGAACATGCAAAAGGTCAT	540	
Qy	541	CAATGCAGCTCAAGTCAGTTGCCCTTTTCTAAGTTTGGAGAAAGCTGTATTCTGTACGGGT	600	
Db	541	CAATGCAGCTCAAGTCAGTTGCCCTTTTCTAAGTTTGGAGAAAGCTGTATTCTGTACGGGT	600	
Qy	601	GGAAGAGATGGCAACATTTATGGTCTGGGATACAGGTGCAACAAAAGATGGGTTTTAT	660	
Db	601	GGAAGAGATGGCAACATTTATGGTCTGGGATACAGGTGCAACAAAAGATGGGTTTTAT	660	
Qy	661	AGGCAAGTGAATCAAAATCAGTGGAGCTCAAAATCCTCAGACAAGCAAAACCCCTTCAAAA	720	
Db	661	AGGCAAGTGAATCAAAATCAGTGGAGCTCAAAATCCTCAGACAAGCAAAACCCCTTCAAAA	720	
Qy	721	CCCAGAGAAACAGAAATTCAAAAGGACTTGCTCTCTTCTGTGGATTTCCAGCAAAAGTGTT	780	
Db	721	CCCAGAGAAACAGAAATTCAAAAGGACTTGCTCTCTTCTGTGGATTTCCAGCAAAAGTGTT	780	
Qy	781	ACTGTGGTCTCTTTCAAGACAGAAATACCTTAGTCTTCAGCAGGAGCTGTGGATGGGATA	840	
Db	781	ACTGTGGTCTCTTTCAAGACAGAAATACCTTAGTCTTCAGCAGGAGCTGTGGATGGGATA	840	

1921 AAGGACTCTCTAGGTCCTACCAATCAAGCAAAATTTGAAGGAGCTGGTACCAAGTATCTCA 1980
1921 AAGGACTCTCTAGGTCCTACCAATCAAGCAAAATTTGAAGGAGCTGGTACCAAGTATCTCA 1980
1981 GAGCCTCCGTCCTCTATCATGTCGGTATGCTTCGAAAGCTGTGGAACGCTACTCTTCTCT 2040
1981 GAGCCTCCGTCCTCTATCATGTCGGTATGCTTCGAAAGCTGTGGAACGCTACTCTTCTCT 2040
2041 TTGAGACCTTTGTGGAAGGCTCTGAAATGGTAGGCAAGAGAAATAGTTCCCCAGAGAAAT 2100
2041 TTGAGACCTTTGTGGAAGGCTCTGAAATGGTAGGCAAGAGAAATAGTTCCCCAGAGAAAT 2100
2101 AAAAATCTGTTGTGTCCTATGCGAGCCAAACGGAAGCTGTGGAATCCATCTCCACGAAGT 2160
2101 AAAAATCTGTTGTGTCCTATGCGAGCCAAACGGAAGCTGTGGAATCCATCTCCACGAAGT 2160
2161 CGGTATCCAGACACCAATTTCCAGGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTC 2220
2161 CGGTATCCAGACACCAATTTCCAGGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTC 2220
2221 ACCATCAGCCCGAGCTCCATGAGGAAATCTGCACATATCTCCATAGAAAGTCCACAGGAG 2280
2221 ACCATCAGCCCGAGCTCCATGAGGAAATCTGCACATATCTCCATAGAAAGTCCACAGGAG 2280
2281 GACTTCTGTGTCCTGAACACTCAACAGAAATATAGATTTCTAATCTGAGTGAGTTACTGA 2340
2281 GACTTCTGTGTCCTGAACACTCAACAGAAATATAGATTTCTAATCTGAGTGAGTTACTGA 2340
2341 GCTTTGTCCTCAATAAAGCTGAGCTTTGTCCTCACTAAACAGATGAAATAACAAG 2400
2341 GCTTTGTCCTCAATAAAGCTGAGCTTTGTCCTCACTAAACAGATGAAATAACAAG 2400
2401 AGTGACTCTATTAATCTGCTGTTTGAAGAGCTGCTTTTCAATTTTGAACAAATCTTT 2460
2401 AGTGACTCTATTAATCTGCTGTTTGAAGAGCTGCTTTTCAATTTTGAACAAATCTTT 2460
2461 TCAACGCTGAAATGTACCTAAATCTGCTTCTACTACCAATATGTATGAGCTTCCCGAG 2520
2461 TCAACGCTGAAATGTACCTAAATCTGCTTCTACTACCAATATGTATGAGCTTCCCGAG 2520
2521 GATGAATGCTGTTTAAATTTTCAATAAAGTAAATTTGTCACTCTAGCAATTTTGAATGAAT 2580
2521 GATGAATGCTGTTTAAATTTTCAATAAAGTAAATTTGTCACTCTAGCAATTTTGAATGAAT 2580
2581 AGTCTTCACTTTTAAATTTTCAATCTCTCTATATATGACATCCCAAGTTCATGGAGG 2640
2581 AGTCTTCACTTTTAAATTTTCAATCTCTCTATATATGACATCCCAAGTTCATGGAGG 2640
2641 CAAAAACAAAGTTTCTTTATCTGTAACCTTTCTATGCTCAGTGGAAGTATCTGCCAG 2700
2641 CAAAAACAAAGTTTCTTTATCTGTAACCTTTCTATGCTCAGTGGAAGTATCTGCCAG 2700
2701 CCACAGATGAGGCTGTGAAGGCTGACTGAGAAATCTCTCTGCTGAAGACCCCTGGTTCT 2760
2701 CCACAGATGAGGCTGTGAAGGCTGACTGAGAAATCTCTCTGCTGAAGACCCCTGGTTCT 2760
2761 GTTCTGCTCCACATGATATTTTATTTGAATACATAAATCTTTTCACTATGAAAAA 2820
2761 GTTCTGCTCCACATGATATTTTATTTGAATACATAAATCTTTTCACTATGAAAAA 2820
2821 AAAAAAATAA 2831
2821 AAAAAAATAA 2831

RESULT 2
LOCUS AX213281 2838 bp DNA linear PAT 06-SEP-2001
DEFINITION Sequence 1 from Patent WO0159115.
ACCESSION AX213281
VERSION AX213281.1 GI:15524191
KEYWORDS Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
REFERENCE
AUTHORS Hubert.R.S., Afar,D.E., Challita-Bid,P.M., Paris,M., Levin,E.,
Mitchell,S.C. and Jakobovits,A.
TITLE 83p5q4: a tissue specific protein highly expressed in prostate
cancer
JOURNAL Patent: WO 0159115-A 1 16-AUG-2001;
UROGENESYS, INC. (US)
FEATURES
source Location/Qualifiers
1..2838
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
130..2322
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC69322.1"
/db_xref="GI:15524192"
/translation="MLFNSALRQPOLGVLBNWSSQYPLQSLLTGYCSGNDHTSYG
ETGVPPVPGCTFSSAPMEHVLAVANEEGFVRLYNTESQSRKCKFKEMWHAHNAVF
DLAWPGELKLVTAAGDQFWDVYKAGELIGTCKHQCSLKSAVAFKPEKAVFTCGG
RDGIMVMDTRCNKDGFRVNOQISGAHNTSDKQTPKPKKQNSKGLAPSDVFOQS
VTVLFDENTLVSAGAVDGIKVMDLKNTAYRQEPISKSPFLYKGSSTPKLGYSS
LILDSTGLTANCTDDNIYFNMGTGLTSVPIFNHONSTFYKSSLSPDQFLVS
GSSDEAYIKVSTPWPQPTVLLGHSQEVTSVCWCPDFTKIATCSDONTLKLMLNR
GLEKPGDGLKLTGWSAKKESRPGLVTVSSQSTPAKAPRVKCNFNSPSSAAAC
APSCAGDLPFNSPTFTSIKTPAKARPINRRSVSVSPKPPSKFMSINRWYTRT
PSSPPIPTPASETKIMSPRLALIPVQKSOAEASERNRKRLSSCUCIESPPVKQ
CVKCNCTVPELSDGOVENLHDLCLLAGNQEDLSGLGPTKSKLIEGAGTISIEPPSP
ISPVAESCGTLPPLRPCGRSEVMGKENSSENKMWLLAMAAKRAKKNPSPSPSS
QTPNSRRSGKTLPSPTVITPSSMRKICTYFHRKSQEDPCGPEHSTEL"

Query Match 99.5%; Score 2816.2; DB 2; Length 2838;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2821; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 3 CACGAGCGGAGTTGGAGCGGATACGATTTGTGTGTGAGAGCGCAACGTGCGATTTTC 62
DB 9 CAGTGGCGGAGTTGGAGCGGATACGATTTGTGTGTGAGAGCGCAACGTGCGATTTTC 68
QY 63 TGCTGAACTTGAGGACATTTCTACGACTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCT 122
DB 69 TGCTGAACTTGAGGACATTTCTACGACTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCT 128
QY 123 GATGCTCTTCAATTCGGTGTCTCCGCGCAGCTTGGCGCTCTGAGAAATGGATGGTC 182
DB 129 GATGCTCTTCAATTCGGCGCTCCGCGCAGCTTGGCGCTCTGAGAAATGGATGGTC 188
QY 183 TTCACAAATACCTTTTCAATCCCTTCTCAGTGGTTATCAGTGCAGTGGTAAATGATGAACA 242
DB 189 TTCACAAATACCTTTTCAATCCCTTCTCAGTGGTTATCAGTGCAGTGGTAAATGATGAACA 248
QY 243 CACTTCTTATGGAGAAACAGGAGTCCGAGTTCCTCTTTTGGATGATGATGATGATGATGAT 302
DB 249 CACTTCTTATGGAGAAACAGGAGTCCGAGTTCCTCTTTTGGATGATGATGATGATGATGAT 308
QY 303 TCCCAATATGGAACATGACTAGCAGTTGCCAATGAAGAGGCTTTGTTTCGATTGTATATA 362
DB 309 TCCCAATATGGAACATGACTAGCAGTTGCCAATGAAGAGGCTTTGTTTCGATTGTATATA 368
QY 363 CACAGAAATCAAAAGTTTTCAGAAAGAGTGTCTTCAAGAAATGGATGGTCTCACTGGAATGC 422
DB 369 CACAGAAATCAAAAGTTTTCAGAAAGAGTGTCTTCAAGAAATGGATGGTCTCACTGGAATGC 428
QY 423 CGTCTTTGACCTGCGCTGGGTTCTGGTGAACCTTAACTTTTACAGCAGCAGGAGTATCA 482
DB 429 CGTCTTTGACCTGCGCTGGGTTCTGGTGAACCTTAACTTTTACAGCAGCAGGAGTATCA 488
QY 483 AACAGCCAAATTTTGGGACGTAAGAGCTGGTGGATTTGGAAACATGCAAAAGGTCATCA 542

ORIGIN

|||||
489 AACGCCAAATTTTGGGACGTAAAGCTGGTGAGCTGATTGGAAATGCAAAAGGTCATCA 548
QY
543 ATGCAGCCTCAAGTCAGTTGCCCTTTTCTAAAGTTTGAGAAAGCTGTATTCTGTACGGGTGG 602
Db
549 ATGCAGCCTCAAGTCAGTTGCCCTTTTCTAAAGTTTGAGAAAGCTGTATTCTGTACGGGTGG 608
QY
603 AAGAGATGGCAACATTATGGTCTGGGATACACAGGTGGCAACAAAAAGATGGGTTTATAG 662
Db
609 AAGAGATGGCAACATTATGGTCTGGGATACACAGGTGGCAACAAAAAGATGGGTTTATAG 668
QY
663 GCAAGTGAATCAAAATCAGTGGAGCTCACAAATACCTCAGACAGCAAAACCCCTTCAAAACC 722
Db
669 GCAAGTGAATCAAAATCAGTGGAGCTCACAAATACCTCAGACAGCAAAACCCCTTCAAAACC 728
QY
723 CAAGAAGAAACAGAAATTCAAAAGGACTTGTCTCCTTCTGTGGATTTCCAGCAAGTGTAC 782
Db
729 CAAGAAGAAACAGAAATTCAAAAGGACTTGTCTCCTTCTGTGGATTTCCAGCAAGTGTAC 788
QY
783 TGTGTCCTCTTTCAAGACGAGAATACCTTAGTCTCAGCAGGAGCTGTGGATGGATTAAT 842
Db
789 TGTGTCCTCTTTCAAGACGAGAATACCTTAGTCTCAGCAGGAGCTGTGGATGGATTAAT 848
QY
843 CAAAGTATGGGATTTACGTAAGAAATTAATCTGCTTATCGACAAGAACCCATAGCATCAA 902
Db
849 CAAAGTATGGGATTTACGTAAGAAATTAATCTGCTTATCGACAAGAACCCATAGCATCAA 908
QY
903 GTCTTTCTCTGTAACCCAGGTAGCAGCACTCGAAAACTTGGATATTCAAAGTCTGATTTTGA 962
Db
909 GTCTTTCTCTGTAACCCAGGTAGCAGCACTCGAAAACTTGGATATTCAAAGTCTGATTTTGA 968
QY
963 TTCCACTGGCTCTACTTTATTTGCTAAATTTGCAAGATTAATGCAAGATTAATCATGTTTAAAT 1022
Db
969 TTCCACTGGCTCTACTTTATTTGCTAAATTTGCAAGATTAATGCAAGATTAATCATGTTTAAAT 1028
QY
1023 GACTGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACCAAGACTCTACCTTTTA 1082
Db
1029 GACTGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACCAAGACTCTACCTTTTA 1088
QY
1083 TGTAAAAATCCAGCCTTAGTCCAGATGACCAAGTTTTTTAGTCAAGTGGCTCAAGTGAAGC 1142
Db
1089 TGTAAAAATCCAGCCTTAGTCCAGATGACCAAGTTTTTTAGTCAAGTGGCTCAAGTGAAGC 1148
QY
1143 TGCCTACATATGGAAGGTTCTCACACCTGGCAACCTCTACTGTGCTCTGGGTCAATTC 1202
Db
1149 TGCCTACATATGGAAGGTTCTCACACCTGGCAACCTCTACTGTGCTCTGGGTCAATTC 1208
QY
1203 TCAAGAGGTCAGTCTGTGCTGGTGTCCATCTGACTTCAAAAAGATTTGCTACCTGTTTC 1262
Db
1209 TCAAGAGGTCAGTCTGTGCTGGTGTCCATCTGACTTCAAAAAGATTTGCTACCTGTTTC 1268
QY
1263 TGATGACAAATACATAAATAATCTGGCGCTTGAATAGAGCTTTAGAGGAAACCCAGGAG 1322
Db
1269 TGATGACAAATACATAAATAATCTGGCGCTTGAATAGAGCTTTAGAGGAAACCCAGGAG 1328
QY
1323 TGATAAATCTTTCCAGGTGGGTTGGGCTCTCAGAAAGAAAAAGAGTCAAGCTTGGCT 1382
Db
1329 TGATAAATCTTTCCAGGTGGGTTGGGCTCTCAGAAAGAAAAAGAGTCAAGCTTGGCT 1388
QY
1383 AGTAACAGTAAACAGTAGCCAGGTACTCTCTGCCAAAGCCCCAGGGTAAAGTGCAATCC 1442
Db
1389 AGTAACAGTAAACAGTAGCCAGGTACTCTCTGCCAAAGCCCCAGGGTAAAGTGCAATCC 1448
QY
1443 ATCCAAATCTTCCCGTCAATCCGAGCTTGTGCCCAAGCTGTGCTGGAGACCTCCCTCT 1502
Db
1449 ATCCAAATCTTCCCGTCAATCCGAGCTTGTGCCCAAGCTGTGCTGGAGACCTCCCTCT 1508
QY
1503 TCCCTTCAAAATACCTCTAGCTTCTATTAAACCTCTCTGCCAAAGCCCGGTCTCCCAT 1562
Db
1509 TCCCTTCAAAATACCTCTAGCTTCTATTAAACCTCTCTGCCAAAGCCCGGTCTCCCAT 1568
QY
1563 CAAACAGAGGCTGTGCTCTCTCGTCTCTCCCAAGCCACCTTCATCTTTTCAAGATGTC 1622
|||||

Db
1569 CAACAGAGAGGCTCTGTCTCTCCGTCTCTCCCAAGCCACCTTTCACTTTTCAAGATGTC 1628
QY
1623 GATTAGAAAATCTGGGTGACCCGAAACACCTTCTCTATCAACCCCATCTACTCCACCTGCTTC 1682
Db
1629 GATTAGAAAATCTGGGTGACCCGAAACACCTTCTCTATCAACCCCATCTACTCCACCTGCTTC 1688
QY
1683 GGACACCAAGATCATGTCTCCGAAAGCCCTTATTCTCTGTGAGCCAGAGTCAATCCCA 1742
Db
1689 GGACACCAAGATCATGTCTCCGAAAGCCCTTATTCTCTGTGAGCCAGAGTCAATCCCA 1748
QY
1743 AGCAGAGCCTTGTCTCTGAGTCTAGAAATAGAGTAAGAGGAGGCTAGACTCAAGCTGTCT 1802
Db
1749 AGCAGAGCCTTGTCTCTGAGTCTAGAAATAGAGTAAGAGGAGGCTAGACTCAAGCTGTCT 1808
QY
1803 GGAGAGTGTGAAAACAAAAGTGTGAAAGTGTGTAACCTGTGTGACTGAGCTTGTATGGCCA 1862
Db
1809 GGAGAGTGTGAAAACAAAAGTGTGAAAGTGTGTAACCTGTGTGACTGAGCTTGTATGGCCA 1868
QY
1863 AGTTGAAAATCTTCAATTTGGATCTGTCTGCTTGTGTAACCCAGGAAGACCTTAGTAA 1922
Db
1869 AGTTGAAAATCTTCAATTTGGATCTGTCTGCTTGTGTAACCCAGGAAGACCTTAGTAA 1928
QY
1923 GGACTCTCTAGTCTTACCAAAATCAAGCAAAATTTGAAGAGCTGTGTACCAAGTATCTCAGA 1982
Db
1929 GGACTCTCTAGTCTTACCAAAATCAAGCAAAATTTGAAGAGCTGTGTACCAAGTATCTCAGA 1988
QY
1983 GCCTCCGTCTCCTATCAGTCCGTATGCTTCAAGAAAGCTGTGGAAACGCTACTCTTCTTT 2042
Db
1989 GCCTCCGTCTCCTATCAGTCCGTATGCTTCAAGAAAGCTGTGGAAACGCTACTCTTCTTT 2048
QY
2043 GAGACCTTGTGGAGAGGCTCTGAAATGGTAGGCAAGAGAAATAGTTCCCCAGAGAAATAA 2102
Db
2049 GAGACCTTGTGGAGAGGCTCTGAAATGGTAGGCAAGAGAAATAGTTCCCCAGAGAAATAA 2108
QY
2103 AAACTGGTGTGTGGCCATGGCAGCAAAACGGAAGGCTGAGAAATCCATCTCCACGAAGTCC 2162
Db
2109 AAACTGGTGTGTGGCCATGGCAGCAAAACGGAAGGCTGAGAAATCCATCTCCACGAAGTCC 2168
QY
2163 GTCAATCCAGACACCCAAATTCAGSAGACAGAGCGGAAAGACATTCGCAAGCCCGGTCAAC 2222
Db
2169 GTCAATCCAGACACCCAAATTCAGSAGACAGAGCGGAAAGACATTCGCAAGCCCGGTCAAC 2228
QY
2223 CATCAAGCCAGCTCCATGAGGAAATCTGCACATCTTCCATAGAAAAGTCCCAAGGAGA 2282
Db
2229 CATCAAGCCAGCTCCATGAGGAAATCTGCACATCTTCCATAGAAAAGTCCCAAGGAGA 2288
QY
2283 CTTCTGTGGTCTCTGAAACACATCAACAGAAATTAAGATTTCTAATCTGAGTGAATTAAGC 2342
Db
2289 CTTCTGTGGTCTCTGAAACACATCAACAGAAATTAAGATTTCTAATCTGAGTGAATTAAGC 2348
QY
2343 TTTGGTCCACTAAAACAAAGCTGAGCTTTGGTCCACTAAAACAAAGTGAATAATCAAGAG 2402
Db
2349 TTTGGTCCACTAAAACAAAGCTGAGCTTTGGTCCACTAAAACAAAGTGAATAATCAAGAG 2408
QY
2403 TGACTCTATAACTCTGGTCTTTAAGAAAGCTGCCCTTTTCAATTTTGAACAAAATCTTTTC 2462
Db
2409 TGACTCTATAACTCTGGTCTTTAAGAAAGCTGCCCTTTTCAATTTTGAACAAAATCTTTTC 2468
QY
2463 AACGCTGAAAATGTACTAATCTGGTCTTACTACCAATATGATATGACGCTTCCCGAGA 2522
Db
2469 AACGCTGAAAATGTACTAATCTGGTCTTACTACCAATATGATATGACGCTTCCCGAGA 2528
QY
2523 TGAATGCTGTGTTTAAATTTTCAATTAAGTAAATTTGTCTACTCTAGCATTTTGAATGAATAG 2582
Db
2529 TGAATGCTGTGTTTAAATTTTCAATTAAGTAAATTTGTCTACTCTAGCATTTTGAATGAATAG 2588
QY
2583 TCTTCACCTTTTAAATTTTCAATTTTCAATTAATGAATCCAGTTCATGAGGAGA 2642
Db
2589 TCTTCACCTTTTAAATTTTCAATTTTCAATTTAATGAATCCAGTTCATGAGGAGA 2648
QY
2643 AAAACCAAGTTTCTTGTGTTATCTCTCAAACTTTCTATGTCTAGTGGAAAGTATCTGCCAGC 2702
Db
2649 AAAACCAAGTTTCTTGTGTTATCTCTCAAACTTTCTATGTCTAGTGGAAAGTATCTGCCAGC 2708
|||||


```
1563 CAACAGAGAGGCTCTGTCTCTCGTCTCTCCCAAGCCACCTTCATCTTCAAGATGTC 1622
1568 CAACAGAGAGGCTCTGTCTCTCGTCTCTCCCAAGCCACCTTCATCTTCAAGATGTC 1627
1623 GATTAGAAACTGGGTGAGCCGAAACACCTTCTCATCACCAACCCATCACTCCACCTGCTTC 1682
1628 GATTAGAAACTGGGTGAGCCGAAACACCTTCTCATCACCAACCCATCACTCCACCTGCTTC 1687
1683 GGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATCTCTGTGAGCCAGAGTCAATCCCA 1742
1688 GGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATCTCTGTGAGCCAGAGTCAATCCCA 1747
1743 AGCAGAGGCTCTCTGAGTCTAGAAATAGAGTAAGAGGAGGCTAGACTCAAGCTGCT 1802
1748 AGCAGAGGCTCTCTGAGTCTAGAAATAGAGTAAGAGGAGGCTAGACTCAAGCTGCT 1807
1803 GGAGAGTGTGAAACAAAGTGTGGAAGTGTAACTGTGTGACTGAGCTTGTATGGCCA 1862
1808 GGAGAGTGTGAAACAAAGTGTGGAAGTGTAACTGTGTGACTGAGCTTGTATGGCCA 1867
1863 AGTTGAAATCTTCAATTTGGATCTGTGCTGCTTGTCTGTTAACTGTAACAGGAGACCTTGTATG 1922
1868 AGTTGAAATCTTCAATTTGGATCTGTGCTGCTTGTCTGTTAACTGTAACAGGAGACCTTGTATG 1927
1923 GGACTCTCTAGTCTCAACCAATCAAGCAAAATTAAGAGAGCTGTGATCACTATCTAGA 1982
1928 GGACTCTCTAGTCTCAACCAATCAAGCAAAATTAAGAGAGCTGTGATCACTATCTAGA 1987
1983 GCCTCCGCTCTCTATCAGTCCGTATGCTTCAGAAAGCTGTGGAAGCTACCTCTTCCCTTT 2042
1988 GCCTCCGCTCTCTATCAGTCCGTATGCTTCAGAAAGCTGTGGAAGCTACCTCTTCCCTTT 2047
2043 GAGACCTTGTGAGAGGCTGTGAAATGGTAGGCAAGAGAAATAGTTCACAGAGAAATG 2102
2048 GAGACCTTGTGAGAGGCTGTGAAATGGTAGGCAAGAGAAATAGTTCACAGAGAAATG 2107
2103 AAACCTGGTGTGGCCATGGCAGCCAAACGAGGCTGAGATCCATCTCCACGAGTCC 2162
2108 AAACCTGGTGTGGCCATGGCAGCCAAACGAGGCTGAGATCCATCTCCACGAGTCC 2167
2163 GTCAATCCAGACACCAATTCAGAGAGCAGAGCGGAAAGACATTCGCAAGCCCGGTCA 2222
2168 GTCAATCCAGACACCAATTCAGAGAGCAGAGCGGAAAGACATTCGCAAGCCCGGTCA 2227
2223 CATCAGCCAGCTCCATGAGGAAATCTGCACATATTCATATAGAAAGTCCAGAGGGA 2282
2228 CATCAGCCAGCTCCATGAGGAAATCTGCACATATTCATATAGAAAGTCCAGAGGGA 2287
2283 CTTCTGTGCTGCTGAACACTCAACAGAAATATAGATTTCTATCTGAGTGAATCTGAGC 2342
2288 CTTCTGTGCTGCTGAACACTCAACAGAAATATAGATTTCTATCTGAGTGAATCTGAGC 2347
2343 TTGTGCTCACTAAACAAAGCTGAGCTTTGGTCCACTAAAAAAGATGAAATACAAAG 2402
2348 TTGTGCTCACTAAACAAAGCTGAGCTTTGGTCCACTAAAAAAGATGAAATACAAAG 2407
2403 TGACTCTATACTGTGCTTTTAAAGAGCTGCCCTTTTCAATTTTATGACAAATCTTTTC 2462
2408 TGACTCTATACTGTGCTTTTAAAGAGCTGCCCTTTTCAATTTTATGACAAATCTTTTC 2467
2463 AACGCTGAAATGATCACTATCTGTTCTACTACCAATATGATATGACGCTTCCGAGGA 2522
2468 AACGCTGAAATGATCACTATCTGTTCTACTACCAATATGATATGACGCTTCCGAGGA 2527
2523 TGAATGCTGTGTTTAAATTTTCAATAAGTAAATTTGCTCACTCTAGCAATTTGAAATGATG 2582
2528 TGAATGCTGTGTTTAAATTTTCAATAAGTAAATTTGCTCACTCTAGCAATTTGAAATGATG 2587
2583 TCTTCACTTTTAAATTTTCAATCTTCTATATAATGACATCCACGTTTATGAGGGA 2642
2588 TCTTCACTTTTAAATTTTCAATCTTCTATATAATGACATCCACGTTTATGAGGGA 2647
```

```
2643 AAAAAACAAGTTTCTTGTATTCTCTGAAACTTTCTATGTCTAGTGAAAGTATCTGCCAGCC 2702
2648 AAAAAACAAGTTTCTTGTATTCTCTGAAACTTTCTATGTCTAGTGAAAGTATCTGCCAGCC 2707
2703 ACAGCATGAGGCTGTGAAAGCTGACTGAGAAATCTCTGCTGTGAGAACCCCTGTTCTGT 2762
2708 ACAGCATGAGGCTGTGAAAGCTGACTGAGAAATCTCTGCTGTGAGAACCCCTGTTCTGT 2767
2763 TCTGCTCCACATGATATAATTTTATTTGAAATACATAATCTTTTCACTATG 2814
2768 TCTGCTCCACATGATATAATTTTATTTGAAATACATAATCTTTTCACTATG 2819

RESULT 4
CQ493832
LOCUS CQ493832 4422 bp DNA linear PAT 30-JAN-2004
DEFINITION Sequence 25699 from Patent WO0160860.
ACCESSION CQ493832
VERSION CQ493832.1 GI:41459451
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Schlegel R., Endege W.O. and Monahan J.E.
TITLE Genes differentially expressed in human prostate cancer and their
use
JOURNAL Patent: WO 0160860-A 25699 23-AUG-2001;
Millennium Predictive Medicine, Inc. (US)
FEATURES
source
1..4422
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.0%; Score 2804; DB 2; Length 4422;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2807; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CACGAGCGGAGTGTGGAGCGATAACGATTTGTGTGTGAGAGCGCAACCTGGCATTTTC 62
DB 8 CAGTGGCGGAGTGTGGAGCGATAACGATTTGTGTGTGAGAGCGCAACCTGGCATTTTC 67
QY 63 TGCTGAACTTGGAGCAATTTCTACACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCT 122
DB 68 TGCTGAACTTGGAGCAATTTCTACGACTTTTCTCAGCTGAGGCTTTTCTCCGACCT 127
QY 123 GATGCTCTTCAATTCGGTGTCTCCGCGAGCCCGAGCTTGGCGTCTGAGAAATGGATGTC 182
DB 128 GATGCTCTTCAATTCGGTGTCTCCGCGAGCCCGAGCTTGGCGTCTGAGAAATGGATGTC 187
QY 183 TTCAATATACCTCTTCAATCCCTTCTGACTGGTATTCAGTGCAAGTGGTATGATGAACA 242
DB 188 TTCAATATACCTCTTCAATCCCTTCTGACTGGTATTCAGTGCAAGTGGTATGATGAACA 247
QY 243 CACTTCTTATGAGAAACAGGAGTCCAGATTCCTCTTTTGGATGTATACCTTCTCTTTCG 302
DB 248 CACTTCTTATGAGAAACAGGAGTCCAGATTCCTCTTTTGGATGTATACCTTCTCTTTCG 307
QY 303 TCCCAATATGAAACATGACTAGCAGTTGCCAATGAAAGAGGCTTTGTTCCGATTGTATAA 362
DB 308 TCCCAATATGAAACATGACTAGCAGTTGCCAATGAAAGAGGCTTTGTTCCGATTGTATAA 367
QY 363 CACAGAAATCAAAAGTTTCAGAAAGAGTGTCTTCAAGAAATGGATGGTCTCACTGGAATGC 422
DB 368 CACAGAAATCAAAAGTTTCAGAAAGAGTGTCTTCAAGAAATGGATGGTCTCACTGGAATGC 427
QY 423 CGTCTTGAACCTGCGCTGGGTTCTGGTGAACTTAAACTTGTTCAGCAGCAGGATGATCA 482
DB 428 CGTCTTGAACCTGCGCTGGGTTCTGGTGAACTTAAACTTGTTCAGCAGCAGGATGATCA 487
```

QY 483 AACAGCCAAATTTTGGACGCTAAAGCTGGTGGAGCTGATTGGAAACATGCAAGGTCATCA 542
DB 488 AACAGCCAAATTTTGGACGCTAAAGCTGGTGGAGCTGATTGGAAACATGCAAGGTCATCA 547
QY 543 ATGAGCCTCAAGTCAGTTGGCTTTTCTAAGTTTGAAGAAAGCTGTATTCTGTACGGGTGG 602
DB 548 ATGAGCCTCAAGTCAGTTGGCTTTTCTAAGTTTGAAGAAAGCTGTATTCTGTACGGGTGG 607
QY 603 AAGAGATGCAACATTTATGGCTGCGGATACAGAGTGCACAAAGATGGGTTTATAG 662
DB 608 AAGAGATGCAACATTTATGGCTGCGGATACAGAGTGCACAAAGATGGGTTTATAG 667
QY 663 GCAAGTGAATCAAAATCAGTGGAGCTCACAAATACCTCAGACAGCAAAACCCCTTCAAAACC 722
DB 668 GCAAGTGAATCAAAATCAGTGGAGCTCACAATACCTCAGACAGCAAAACCCCTTCAAAACC 727
QY 723 CAAGAAGAAACAGAAATTCAAAAGGACTTGCTCCTTCTGTGGATTTCCAGCAAGGTGTAC 782
DB 728 CAAGAAGAAACAGAAATTCAAAAGGACTTGCTCCTTCTGTGGATTTCCAGCAAGGTGTAC 787
QY 783 TGTGCTCCTTTTCAAGCAGAAATACCTTAGTCTCAGCAGAGCTGTGGATGGAATAT 842
DB 788 TGTGCTCCTTTTCAAGCAGAAATACCTTAGTCTCAGCAGAGCTGTGGATGGAATAT 847
QY 843 CAAAGTATGGGATTTACGTAAGAAATTAATCTGCTTATCGACAAGAACCCATAGCATCCAA 902
DB 848 CAAAGTATGGGATTTACGTAAGAAATTAATCTGCTTATCGACAAGAACCCATAGCATCCAA 907
QY 903 GTCTTTCTGTACCCAGTACGAGCAGTACGAAATCTTGGATATCAAGTCTGATTTTGA 962
DB 908 GTCTTTCTGTACCCAGTACGAGCAGTACGAAATCTTGGATATCAAGTCTGATTTTGA 967
QY 963 TTCACCTGGCTCTACTTTATTTGCTAATTTGACAGACGATACATCTACATGTTTAAAT 1022
DB 968 TTCACCTGGCTCTACTTTATTTGCTAATTTGACAGACGATACATCTACATGTTTAAAT 1027
QY 1023 GACTGGTGTGAAGACTTCTCAGTGGCTATTTTCAATGGACACCAAGACTCTACCTTTTA 1082
DB 1028 GACTGGTGTGAAGACTTCTCAGTGGCTATTTTCAATGGACACCAAGACTCTACCTTTTA 1087
QY 1083 TGTAAATCCAGCCTTAGTCCAGATGACAGTGTGTTTGTAGTGGCTCAAGTGTGAAGC 1142
DB 1088 TGTAAATCCAGCCTTAGTCCAGATGACAGTGTGTTTGTAGTGGCTCAAGTGTGAAGC 1147
QY 1143 TGCCTACATATGGAAGTCTCACAACCTGGCAACCTCTACTGTGCTCTGGGTCAATTC 1202
DB 1148 TGCCTACATATGGAAGTCTCACAACCTGGCAACCTCTACTGTGCTCTGGGTCAATTC 1207
QY 1203 TCAAGAGTTCAGTCTGTGCTGTGCTGCTCCTACATCTCACAAGATTTGCTACCTGTTTC 1262
DB 1208 TCAAGAGTTCAGTCTGTGCTGTGCTGCTCCTACATCTCACAAGATTTGCTACCTGTTTC 1267
QY 1263 TGATGACAATACATAAATAATCTGGCGCTTGAATAGAGCTTAGAGGAAACACAGGAGG 1322
DB 1268 TGATGACAATACATAAATAATCTGGCGCTTGAATAGAGCTTAGAGGAAACACAGGAGG 1327
QY 1323 TGATAAATCTTCCAGTGGGTGGGCTCTCAGAAGAAATAAGAGTCAAGACCTGGGCT 1382
DB 1328 TGATAAATCTTCCAGTGGGTGGGCTCTCAGAAGAAATAAGAGTCAAGACCTGGGCT 1387
QY 1383 AGTAACAGTAAACAGTACGAGTACTCTGCGCAAGGCCCGGAGTAAAGTGCAATCC 1442
DB 1388 AGTAACAGTAAACAGTACGAGTACTCTGCGCAAGGCCCGGAGTAAAGTGCAATCC 1447
QY 1443 ATCCAAATCTTCCCGCTCATCCGAGCTTGTGCCCAAGCTGTGCTGGAGACCTCCCTCT 1502
DB 1448 ATCCAAATCTTCCCGCTCATCCGAGCTTGTGCCCAAGCTGTGCTGGAGACCTCCCTCT 1507
QY 1503 TCCTTCAAATACCTCTACGTTCTTATTAATAAATCTCTCTGCGCAAGGCCCGGCTCTCCAT 1562
DB 1508 TCCTTCAAATACCTCTACGTTCTTATTAATAAATCTCTCTGCGCAAGGCCCGGCTCTCCAT 1567

QY 1563 CAACAGAGAGGCTGTGCTCTCTCCGTCTCTCCAGCCACCTTTCATCTTTCAAGATGTC 1622
DB 1568 CAACAGAGAGGCTGTGCTCTCTCCGTCTCTCCAGCCACCTTTCATCTTTCAAGATGTC 1627
QY 1623 GATTAGAAACTGGGTGAGCCCGAAACACCTTCTCATACACACCCATCACTCCACCTGCTTC 1682
DB 1628 GATTAGAAACTGGGTGAGCCCGAAACACCTTCTCATACACCCATCACTCCACCTGCTTC 1687
QY 1683 GGAGACCAAGATCATGTCTCGAGAAAGCCCTTATTCCTGTGAGCCAGAGTCAATCCA 1742
DB 1688 GGAGACCAAGATCATGTCTCGAGAAAGCCCTTATTCCTGTGAGCCAGAGTCAATCCA 1747
QY 1743 AGCAGAGGCTTGTCTGTGAGCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCT 1802
DB 1748 AGCAGAGGCTTGTCTGTGAGCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCT 1807
QY 1803 GGAGAGTGTGAAACAAAGTGTGTGAAGATTTGTAACTGTGTGACTGAGCTTGTAGGGCA 1862
DB 1808 GGAGAGTGTGAAACAAAGTGTGTGAAGATTTGTAACTGTGTGACTGAGCTTGTAGGGCA 1867
QY 1863 AGTTGAAATCTTCAATTTGGATCTGTGCTGCTTGTGTTAACCCAGGAAGACCTTAGTAA 1922
DB 1868 AGTTGAAATCTTCAATTTGGATCTGTGCTGCTTGTGTTAACCCAGGAAGACCTTAGTAA 1927
QY 1923 GGACTCTTAGGCTCTACCAATCAAGCAAAATTTGAAGAGCTGTGTACCAAGTATCTCAGA 1982
DB 1928 GGACTCTTAGGCTCTACCAATCAAGCAAAATTTGAAGAGCTGTGTACCAAGTATCTCAGA 1987
QY 1983 GCCTCCGTCTCCTATCAGTCCGTATGTTTCAAGAGCTGTGGAAAGCTTACCTCTTCTTTT 2042
DB 1988 GCCTCCGTCTCCTATCAGTCCGTATGTTTCAAGAGCTGTGGAAAGCTTACCTCTTCTTTT 2047
QY 2043 GAGACCTTGTGAGAGGCTCTGAAATGTTAGAGCTGTGGAAAGCTTACCTCTTCTTTT 2102
DB 2048 GAGACCTTGTGAGAGGCTCTGAAATGTTAGAGCTGTGGAAAGCTTACCTCTTCTTTT 2107
QY 2103 AAACCTGGTGTGTGGCCATGGCAGGCAAAACGGAAGCTGTGAGAAATCCATCTCCACGAAGTCC 2162
DB 2108 AAACCTGGTGTGTGGCCATGGCAGGCAAAACGGAAGCTGTGAGAAATCCATCTCCACGAAGTCC 2167
QY 2163 GTCATCCAGACACCCAAATTCAGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTCAAC 2222
DB 2168 GTCATCCAGACACCCAAATTCAGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTCAAC 2227
QY 2223 CATCACGCCAGCTCCATGAGGAAATCTGCACATCTTCCATAGAAAGTCCCAAGGAGGA 2282
DB 2228 CATCACGCCAGCTCCATGAGGAAATCTGCACATCTTCCATAGAAAGTCCCAAGGAGGA 2287
QY 2283 CTTCTGTGGTCTCTGAACACTCAACAGAAATTTAGATTTCTAATCTGAGTGAATCTGAGC 2342
DB 2288 CTTCTGTGGTCTCTGAACACTCAACAGAAATTTAGATTTCTAATCTGAGTGAATCTGAGC 2347
QY 2343 TTTGGTCCACTTAAACAAAGCTGAGCTTTGGTCCACTAAACCAAGATGAAATCAAGAG 2402
DB 2348 TTTGGTCCACTTAAACAAAGCTGAGCTTTGGTCCACTAAACCAAGATGAAATCAAGAG 2407
QY 2403 TGACTCTATAACTCTGTGCTTTTGAAGAGCTGCTTTTCAATTTTAGCAAAATCTTTTC 2462
DB 2408 TGACTCTATAACTCTGTGCTTTTGAAGAGCTGCTTTTCAATTTTAGCAAAATCTTTTC 2467
QY 2463 AACGCTGAAATGTACCTAATCTGGTCTTACTACCAATATATATGATGAGCTTCCCGAGGA 2522
DB 2468 AACGCTGAAATGTACCTAATCTGGTCTTACTACCAATATATGATGAGCTTCCCGAGGA 2527
QY 2523 TGAATGCTGTGTTTAAATTTTCAAAAGTAAATTTGTCTAGCATTTTGAATGAATAG 2582
DB 2528 TGAATGCTGTGTTTAAATTTTCAAAAGTAAATTTGTCTAGCATTTTGAATGAATAG 2587
QY 2583 TCTTCACTTTTAAATTTTATTCATCTTCTATATATATGACATCCAGTTCATGAGGCA 2642
DB 2588 TCTTCACTTTTAAATTTTATTCATCTTCTATATATATGACATCCAGTTCATGAGGCA 2647
QY 2643 AAAACCAAGTTTCTTGTGTTATCTCTGAAACCTTTTCTATGCTCAGTGGAAAGTATCTGCCAGCC 2702

```
|||||
2648 AAAAAAAGATTCTTGTATCTCTGAAACTTTCTATGCTCAGTGGAAAGTATCTGCCACC 2707
QY
2703 ACAGCATGAGGCTCTGGAAGCTGACTGAGAAATCCTCTGCTGAAGACCCCTGGTTCTGT 2762
Db
2708 ACAGCATGAGGCTCTGGAAGCTGACTGAGAAATCCTCTGCTGAAGACCCCTGGTTCTGT 2767
QY
2763 TCTGCTCAACATGTATAATTTTATTTGAAATACATAATCTTTTCACTATG 2814
Db
2768 TCTGCTCAACATGTATAATTTTATTTGAAATACATAATCTTTTCACTATG 2819

RESULT 5
AX188263
LOCUS AX188263 4422 bp DNA linear PAT 06-AUG-2001
DEFINITION Sequence 3958 from Patent WO0142467.
ACCESSION AX188263
VERSION AX188263.1 GI:15139736
KEYWORDS
SOURCE Homo sapiens (human)
% ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
REFERENCE
AUTHORS Schlegel, R., Deeds, J., Berger, A. and Zhao, X.
TITLE Genes, compositions, kits, and methods for identification,
assessment, prevention, and therapy of cervical cancer
JOURNAL Patent: WO 0142467-A 3958 14-JUN-2001;
Millennium Predictive Medicine, Inc. (US)
FEATURES
source
1..4422
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 99.0%; Score 2804; DB 2; Length 4422;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2807; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CACGAGCGGAGTTCGAGCGATACGATTTGTGTGAGAGCGCAACGTGCGATTTC 62
Db 8 CAGTGGCGGAGTTGGAGCGATAACGATTTGTGTGAGAGCGCAAGCTGCCGATTTC 67
QY 63 TGCTGAACCTTGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCT 122
Db 68 TGCTGAACCTTGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCT 127
QY 123 GATGCTCTTCAATTCGGTCTCCGCCAGCCCGAGCTTGGCGTCTCAGAAATGGATGTC 182
Db 128 GATGCTCTTCAATTCGGTCTCCGCCAGCCCGAGCTTGGCGTCTCAGAAATGGATGTC 187
QY 183 TTCAAAATACCCCTTTCAATCCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACA 242
Db 188 TTCAAAATACCCCTTTCAATCCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACA 247
QY 243 CACTTCTTATGAGAAACAGGAGTCCCAAGTTCCTCTTTTGGATGATACCTTCTCTTCTGC 302
Db 248 CACTTCTTATGAGAAACAGGAGTCCCAAGTTCCTCTTTTGGATGATACCTTCTCTTCTGC 307
QY 303 TCCCAATATGGAACATGTACTAGCAGTTGGCCAAATGAAGAGCTTTGTTGATTTGATATAA 362
Db 308 TCCCAATATGGAACATGTACTAGCAGTTGGCCAAATGAAGAGCTTTGTTGATTTGATATAA 367
QY 363 CACAGAATCACAAAGTTTCAGAAAGAGTGCTTCAAGAAATGGATGGCTCACTGGAATGC 422
Db 368 CACAGAATCACAAAGTTTCAGAAAGAGTGCTTCAAGAAATGGATGGCTCACTGGAATGC 427
QY 423 CGTCTTTGACCTGGCTGGGTTCTGGTGAACCTTAAACCTTTGTTACAGCAGCAGGTGATCA 482
Db 428 CGTCTTTGACCTGGCTGGGTTCTGGTGAACCTTAAACCTTTGTTACAGCAGCAGGTGATCA 487
```

```
QY 483 AACAGCCAAATTTTGGGACGTAAGAGCTGGTGAAGCTGATTTGGAACATGCAAAAGGTCAATCA 542
Db 488 AACAGCCAAATTTTGGGACGTAAGAGCTGGTGAAGCTGATTTGGAACATGCAAAAGGTCAATCA 547
QY 543 ATGCAAGCTCAAGTCAAGTTCGCTTTTCTAAAGTTTGGAGAAAGCTGTATTTCTGTACGGGTGG 602
Db 548 ATGCAAGCTCAAGTCAAGTTCGCTTTTCTAAAGTTTGGAGAAAGCTGTATTTCTGTACGGGTGG 607
QY 603 AAGAGATGGCAACATTTATGCTCTGGGATACAGAGTGCAACAAAGAAAGATGGGTTTATAG 662
Db 608 AAGAGATGGCAACATTTATGCTCTGGGATACAGAGTGCAACAAAGAAAGATGGGTTTATAG 667
QY 663 GCAAGTGAATCAAAATCAGTGGAGCTCAAAATACCTCAGACAGCAAAACCCCTTCAAAACC 722
Db 668 GCAAGTGAATCAAAATCAGTGGAGCTCAAAATACCTCAGACAGCAAAACCCCTTCAAAACC 727
QY 723 CAAGAAGAAACAGAAATTTCAAAAGGACTTGTCTCTTCTGTGGATTTCCAGCAAAAGTGTTC 782
Db 728 CAAGAAGAAACAGAAATTTCAAAAGGACTTGTCTCTTCTGTGGATTTCCAGCAAAAGTGTTC 787
QY 783 TGTGCTCTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGATAT 842
Db 788 TGTGCTCTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGATAT 847
QY 843 CAAAGTATGGGATTTACGTAAGAAATATACCTTATCGACAAAGAACCCATAGCATCCAA 902
Db 848 CAAAGTATGGGATTTACGTAAGAAATATACCTTATCGACAAAGAACCCATAGCATCCAA 907
QY 903 GTCTTTCTCTGTACCCAGGTAGCAGCACTCGAAATCTTGGATATTTCAAGTCTGATTTTGA 962
Db 908 GTCTTTCTCTGTACCCAGGTAGCAGCACTCGAAATCTTGGATATTTCAAGTCTGATTTTGA 967
QY 963 TTCACTGGCTCTACTTTTATTTGCTAAATTTGTCACAGAGATTAACATCTACATGTTTATAT 1022
Db 968 TTCACTGGCTCTACTTTTATTTGCTAAATTTGTCACAGAGATTAACATCTACATGTTTATAT 1027
QY 1023 GACTGGGTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTA 1082
Db 1028 GACTGGGTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTA 1087
QY 1083 TGTAAATTCAGACCTTAGTCCAGATGACCAAGTCTTGTAGTGGCTCAAGTGTATGAGC 1142
Db 1088 TGTAAATTCAGACCTTAGTCCAGATGACCAAGTCTTGTAGTGGCTCAAGTGTATGAGC 1147
QY 1143 TGCCTACATATGGAAGGTCTCCACACCTGGCAACCTCTACTGTCTCTGGGTCAATC 1202
Db 1148 TGCCTACATATGGAAGGTCTCCACACCTGGCAACCTCTACTGTCTCTGGGTCAATC 1207
QY 1203 TCAAGAGTCAAGTCTGTGTGCTGGTCCATCTGACTTTCACAAAGATTTGCTACTCTGTC 1262
Db 1208 TCAAGAGTCAAGTCTGTGTGCTGGTCCATCTGACTTTCACAAAGATTTGCTACTCTGTC 1267
QY 1263 TGATGACAAATACACTAAATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACACGAGG 1322
Db 1268 TGATGACAAATACACTAAATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACACGAGG 1327
QY 1323 TGATAAACTTTCCACGGTGGGTTCGGCTCTCAGAGAAAGAAAGAGTCAAGACCTGGCT 1382
Db 1328 TGATAAACTTTCCACGGTGGGTTCGGCTCTCAGAGAAAGAAAGAGTCAAGACCTGGCT 1387
QY 1383 AGTAACAGTAAACAGTAGCCAGAGTACTCTGTCGCAAGCCCGAGGGTAAAGTGCATCC 1442
Db 1388 AGTAACAGTAAACAGTAGCCAGAGTACTCTGTCGCAAGCCCGAGGGTAAAGTGCATCC 1447
QY 1443 ATCCAAATTTTCCCGCTCATCCGAGCTTGTGCCCGCAAGCTGTCTGGAGACCTCCCTCT 1502
Db 1448 ATCCAAATTTTCCCGCTCATCCGAGCTTGTGCCCGCAAGCTGTCTGGAGACCTCCCTCT 1507
QY 1503 TCCTTCAAAATCTCCTAGTCTCTATTAATAACCTCTCTCGCAAGCCCGGTCTCCCAT 1562
Db 1508 TCCTTCAAAATCTCCTAGTCTCTATTAATAACCTCTCTCGCAAGCCCGGTCTCCCAT 1567
QY 1563 CAACAGAGAGGCTCTGTCTCTCCGCTCTCTCCCAAGCCACCTTCAATCTTTCAAGATGTC 1622
```

1568 CAACGAGAGGCTCTGTCTCTCTCGTCTCTCCAGCCACCTTCATCTTTCAAGATGTC 1627
1623 GATTAGAACTGGTGACCCGGAACACCTTCCTCATCACACCACATCACTCCACCTGCTTC 1682
1628 GATTAGAACTGGTGACCCGGAACACCTTCCTCATCACACCACATCACTCCACCTGCTTC 1687
1683 GGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTCGTGTGAGCCAGAGTCATCCCA 1742
1688 GGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTCGTGTGAGCCAGAGTCATCCCA 1747
1743 AGCAGAGGCTTCTGTAGCTCTAGAAATAGAGTAAAGAGAGGCTAGACTCAAGCTGTCT 1802
1748 AGCAGAGGCTTCTGTAGCTCTAGAAATAGAGTAAAGAGAGGCTAGACTCAAGCTGTCT 1807
1803 GGAGAGTGAAACAAACAGTGTGTGAAGAGTGTAACTGTGTGACTGAGCTTGTATGGCCA 1862
1808 GGAGAGTGTAACAAACAGTGTGTGAAGAGTGTAACTGTGTGACTGAGCTTGTATGGCCA 1867
1863 AGTTGAAATCTTATTTGGATCTGTGCTGCTTGTGTGTAACAGGAGAGACTTAGTAA 1922
1868 AGTTGAAATCTTATTTGGATCTGTGCTGCTTGTGTGTAACAGGAGAGACTTAGTAA 1927
1923 GGACTCTTAGGCTCTACCAATCAAGCAAAATTTGAAGGAGCTGTGTACAGTATCTCAGA 1982
1928 GGACTCTTAGGCTCTACCAATCAAGCAAAATTTGAAGGAGCTGTGTACAGTATCTCAGA 1987
1983 GCTCGCTCTCTATCAGTCCGTATGCTTCAGAAAGCTGTGAAACGCTACCTCTCTCTTT 2042
1988 GCTCGCTCTCTATCAGTCCGTATGCTTCAGAAAGCTGTGAAACGCTACCTCTCTCTTT 2047
2043 GAGACCTCTGTGAGAGGCTCTGAATGTGTGCAAGAGAGATAGTTCCCCAGAGAAATA 2102
2048 GAGACCTCTGTGAGAGGCTCTGAATGTGTGCAAGAGAGATAGTTCCCCAGAGAAATA 2107
2103 AAACCTGTTGTTGGCCATCGCCAAACGGAAGGCTGAGAAATCCATCTCCACGAGTCC 2162
2108 AAACCTGTTGTTGGCCATCGCCAAACGGAAGGCTGAGAAATCCATCTCCACGAGTCC 2167
2163 GTATCCCAAGACCAATTCAGGAGAGACAGAGCGGAAAGATGTCGCAAGCCCGGTAC 2222
2168 GTATCCCAAGACCAATTCAGGAGAGACAGAGCGGAAAGATGTCGCAAGCCCGGTAC 2227
2223 CATCAGCCAGCTCCATCAGAGAAATCTGCACATCTCCATAGAAAGTCCAGGAGGA 2282
2228 CATCAGCCAGCTCCATCAGAGAAATCTGCACATCTCCATAGAAAGTCCAGGAGGA 2287
2283 CTTCTGTGCTCTGAACTCAACAGAAATATAGATTTCTAATCTGAGTGAGTTACTGAGC 2342
2288 CTTCTGTGCTCTGAACTCAACAGAAATATAGATTTCTAATCTGAGTGAGTTACTGAGC 2347
2343 TTTGTGCTCAATAACAGCTGAGCTTGGTCCATAAAACAGATGAAATAACAAG 2402
2348 TTTGTGCTCAATAACAGCTGAGCTTGGTCCATAAAACAGATGAAATAACAAG 2407
2403 TGACTCTAATCTGCTTTTGAAGAGCTGCTTTTCAATTTTATAGAAAATCTTTTC 2462
2408 TGACTCTAATCTGCTTTTGAAGAGCTGCTTTTCAATTTTATAGAAAATCTTTTC 2467
2463 AACGCTGAAATGTACCTAATCTGTTTCTACTACCAATAATATATGACGCTTCCCGAGGA 2522
2468 AACGCTGAAATGTACCTAATCTGTTTCTACTACCAATAATATATGACGCTTCCCGAGGA 2527
2523 TGAATGCTGTTTAAATTTCAATAAGTAAATTTGTCACTCTAGCAATTTGAATGAATAG 2582
2528 TGAATGCTGTTTAAATTTCAATAAGTAAATTTGTCACTCTAGCAATTTGAATGAATAG 2587
2583 TCTTCACTTTTAAATTTCAATCTCTCTAATAATGACATCCAGTTCATGAGGCA 2642
2588 TCTTCACTTTTAAATTTCAATCTCTCTAATAATGACATCCAGTTCATGAGGCA 2647
2643 AAAACAAGTCTTCTGTTTATCTGAACTTTCTATGCTCAGTGGAAAGTATCTGCCAGCC 2702

2648 AAAACAAGTCTTCTGTTTATCTCTGAAACTTTTCTATGTCTCAGTGGAAAGATCTGCCAGCC 2707
2703 ACAGCATGAGGCTGTGAGGCTGACTGAGAAATCTCTGCTGAGAGCCCTGTTCTGT 2762
2708 ACAGCATGAGGCTGTGAGGCTGACTGAGAAATCTCTGCTGAGAGCCCTGTTCTGT 2767
2763 TCTGCTCTCAACATGATATAATTTTATTTTGAATAATATAATCTTTTCACTATG 2814
2768 TCTGCTCTCAACATGATATAATTTTATTTTGAATAATATAATCTTTTCACTATG 2819

RESULT 6
AK000742
LOCUS Homo sapiens cDNA FLJ20735 fis, clone HEP08479.
DEFINITION AK000742
ACCESSION AK000742
VERSION AK000742.1 GI:7021020
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.

REFERENCE 1
AUTHORS Kawakami,T., Nouchi,S., Itoh,T., Shigeta,K., Senba,T.,
Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Ota,T.,
Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
2 (bases 1 to 3008)
AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission
Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1,
Minato-ku, Tokyo 108-8639, Japan (E-mail: flicdna@ims.u-tokyo.ac.jp,
Tel: 81-3-5449-5286, Fax: 81-3-5449-5416)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- and 3'-end one pass sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).

FEATURES
source
1..3008
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="HEP08479"
/cell_lines="HepG2"
/cell_type="hepatoma"
/clone_lib="HEP"
/note="Cloning vector pME18SFL3"
157..2349
/note="unnamed protein product"
/codon_start=1
/protein_id="BAA91355.1"
/trnslation="MLFNSVLRLQPLGLVLRNGWSQYPLQSLITGYQCSGNDEHTSYG
ETGVVPVPPGCTFSSAPNMEHLAVANEEGFVRLYNTEPQSRKKCKPEMAHWNVAF
DLAWPGLKLTWAGDQTAKFWDVKAGELIGTCTGHCQSLKSNVAFKPEKAVFTGG
RDGNIMWDRKNDKGFYQVQNIQGAHNTSDKQTPSKPKKKQNSKGLASVDFQQS
VTYVLFODENTLVAGAVDGIKIVMDLRKNYTAHQEPYASKFLYPCSSSTRKLGYS
LILDSGTSLFANCTDDNIYMFNMTGLKTSFPAIFNGHNSFTFYKSSLSDDQPLVS
GSSDEAAVIWKVSTPWPQPTVLLGHOSVTSVCWCPSTFTKIATCDDNTLKIWLNR
GLEKPGCKLSTVWASQKKESRPGIYVTSQSTPAKAPRVKCNPSNSPSSAAC
APSCAGDULPLSPNTSFTSKTSPAKAPINRRGSVSVPKPPSSFKMSIRNWRTY
PSSPPPTTPPASETKIMSPKRALIPVQSKSQACSESRRNRVRRRLDSSCLSEVKQK
CVKSCNCTVELDQVENLHLDLCLLQAGNEDLSKDSLGTPTKSKIEGATGTSISEPSP
ISPYASBSGCTPLPLRPGGSEMVKGKNSPENKWLMLAARAKRKAENPSRSPSS
QTPNSRQSGKTLPSVTLTTPSSMRKICTYFHRKSKQEDFCGPEHSTL"

ORIGIN																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
--------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

QY 2103 AAATCGTGTGGTGGCCATGCGACGCAAAACGAGGCTGAGATCCATCTCCACGAAGTCC 2162
Db 2136 AAATCGTGTGGTGGCCATGCGACGCAAAACGAGGCTGAGATCCATCTCCACGAAGTCC 2195
QY 2163 GTCATCCGACACCCCAATTCGAGGAGACAGAGCGGAAGACATTCGCCAAGCCCGGTAC 2222
Db 2196 GTCATCCGACACCCCAATTCGAGGAGACAGAGCGGAAGACATTCGCCAAGCCCGGTAC 2255
QY 2223 CATCAGCCCGAGTCCATGAGGAAATCTGCACATATCTCCATAGAAAGTCCCGAGGAG 2282
Db 2256 CATCAGCCCGAGTCCATGAGGAAATCTGCACATATCTCCATAGAAAGTCCCGAGGAG 2315
QY 2283 CTTCTGTGCTCCTGAACACTCAACAGATATAGATTCTTAATCTGAGTGAGTTACTGAGC 2342
Db 2316 CTTCTGTGCTCCTGAACACTCAACAGATATAGATTCTTAATCTGAGTGAGTTACTGAGC 2375
QY 2343 TTGTGTCATTAACAAAGCTGAGCTTTGGTTCACATAAAGAGATGAAAATACAGAG 2402
Db 2376 TTGTGTCATTAACAAAGCTGAGCTTTGGTTCACATAAAGAGATGAAAATACAGAG 2435
QY 2403 TGACTCTATAACTCTGCTTTTAAAGAAAGCTGCTTTTCAATTTTACACAAATCTTTTC 2462
Db 2436 TGACTCTATAACTCTGCTTTTAAAGAAAGCTGCTTTTCAATTTTACACAAATCTTTTC 2495
QY 2463 AACGCTGAAATGTACCTAATCTGTTTCTACTACCATATATATGACAGTTCCTCCGAGG 2522
Db 2496 AACGCTGAAATGTACCTAATCTGTTTCTACTACCATATATATGACAGTTCCTCCGAGG 2555
QY 2523 TGAATGCTGTGTTTAAATTTTCAATAAGTAAATTTGTCACCTAGCATTTTGAATGAATAG 2582
Db 2556 TGAATGCTGTGTTTAAATTTTCAATAAGTAAATTTGTCACCTAGCATTTTGAATGAATAG 2615
QY 2583 TCTTCACTTTTAAATTTTCAATCTCTCTAATAATGACATCCAGTTCATGAGGCA 2642
Db 2616 TCTTCACTTTTAAATTTTCAATCTCTCTAATAATGACATCCAGTTCATGAGGCA 2675
QY 2643 AAAACAAGTCTTCTGTTTATCTCTGAACTTTCTATGCTCAGTGGAAAGTATCTGCAGGC 2702
Db 2676 AAAACAAGTCTTCTGTTTATCTCTGAACTTTCTATGCTCAGTGGAAAGTATCTGCAGGC 2735
QY 2703 ACAGCATGAGGCTGTGAAGGCTGACTGAGAAATCTCTGCTGAGAACCCCTGGTTCGT 2762
Db 2736 ACAGCATGAGGCTGTGAAGGCTGACTGAGAAATCTCTGCTGAGAACCCCTGGTTCGT 2795
QY 2763 TCTGCTCCCAATGATATATTTTATTTTCAATATACATATCTTTTCACTATG 2814
Db 2796 TCTGCTCCCAATGATATATTTTATTTGAAATACATAATCTTTTCACTATG 2847

RESULT 7
BC033540
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BC033540
Homo sapiens denticleless homolog (Drosophila), mRNA (cdna clone
MGC:42657 IMAGE:4826434), complete cds.
BC033540
BC033540.1 GI:23722356
MGC.
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 2969)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altshuler,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diachenko,L., Marusina,K., Farmer,A.A., Rubin,G.W., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L.,
Schetz,T.B., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaby,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,

Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madao,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madao,A., Young,A.C., Shervenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalish,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2969)
NIH MGC Project
Direct Submission
Submitted (26-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 71 Row: 0 Column: 18.
Location/Qualifiers
1..2969
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:42657 IMAGE:4826434"
/tissue_type="Testis"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Vector: pBluescriptR"
1..2969
/gene="DTL"
/note="synonyms: L2DTL, RAMP"
/db_xref="GeneID:51514"
121..2313
/gene="DTL"
/codon_start=1
/product="denticleless homolog"
/protein_id="AAH33540.1"
/db_xref="GI:23722357"
/db_xref="GeneID:51514"

FEATURES
source
gene
CDS
ORIGIN

Query Match		99.0%;	Score 2802.2;	DB 5;	Length 2969;		
Best Local Similarity		99.9%;	Pred. No. 0;				
Matches 2804;		Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;		
Qy	8	GGGGAGTTGGAGGCGATACGATTGTGTCTGTGAGAGCGCAACGTCGGATTTCTGTGTG	67				
Db	5	GGGGAGTTGGAGGCGATACGATTGTGTGTGAGAGCGCAACGTCGGATTTCTGTGTG	64				
Qy	68	AATCTTGGAGGCATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGC	127				
Db	65	AATCTTGGAGGCATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGC	124				
Qy	128	TCTTCAATTCGGTGTCTCCGACGCCAGCTTGGGCTCTCTGAGAAATGATGGTCTTTCAC	187				
Db	125	TCTTCAATTCGGTGTCTCCGACGCCAGCTTGGGCTCTCTGAGAAATGATGGTCTTTCAC	184				
Qy	188	AATACCCCTCTCAATCCCTTCTGACTGGTATACAGTGCAAGTGTATGATGAACACACTT	247				
Db	185	AATACCCCTCTCAATCCCTTCTGACTGGTATACAGTGCAAGTGTATGATGAACACACTT	244				
Qy	248	CTTATGGAGAAACAGAGTCCAGTTTCTCTTTTGGATGTACTCTTCTCTGCTCCCA	307				
Db	245	CTTATGGAGAAACAGAGTCCAGTTTCTCTTTTGGATGTACTCTTCTCTGCTCCCA	304				
Qy	308	ATATGGAACATGTACTAGCAGTTGCCAATGAAGAAGGCTTTGTTTCGATGTATACACAG	367				
Db	305	ATATGGAACATGTACTAGCAGTTGCCAATGAAGAAGGCTTTGTTTCGATGTATACACAG	364				
Qy	368	AATCAAAAGTTTCAGAAAGAGTGCTTCAAGAATGGATGGCTCACTGGAATGCCGTCT	427				
Db	365	AATCAAAAGTTTCAGAAAGAGTGCTTCAAGAATGGATGGCTCACTGGAATGCCGTCT	424				
Qy	428	TTGACCTGSCCTGGGTTCTCGTGAACTTAAACTTTGTACAGCAGCAGGTGATCAACAG	487				
Db	425	TTGACCTGSCCTGGGTTCTCGTGAACTTAAACTTTGTACAGCAGCAGGTGATCAACAG	484				
Qy	488	CCAAATTTTGGAGCGTAAAGCTGTGAGCTGATTTGGAACATGCAAGGTCATCAATGCA	547				
Db	485	CCAAATTTTGGAGCGTAAAGCTGTGAGCTGATTTGGAACATGCAAGGTCATCAATGCA	544				
Qy	548	GCCTCAAGTCAGTGTCTTTCTTAAGTTTGAGAAAGCTGTATTTCTGTCGGGTGGAAG	607				
Db	545	GCCTCAAGTCAGTGTCTTTCTTAAGTTTGAGAAAGCTGTATTTCTGTCGGGTGGAAG	604				
Qy	608	ATGGCAACATTTATGCTCTGGGATACAGGTGCAACAAAGAGATGGTTTTATAGGCAAG	667				
Db	605	ATGGCAACATTTATGCTCTGGGATACAGGTGCAACAAAGAGATGGTTTTATAGGCAAG	664				
Qy	668	TGAATCAAAATCAGTGGAGCTCAATATCCTCAGACAGCAAAACCCCTTCAAAACCCAAAG	727				
Db	665	TGAATCAAAATCAGTGGAGCTCAATATCCTCAGACAGCAAAACCCCTTCAAAACCCAAAG	724				
Qy	728	AGAAACAGAAATTCAGAGGACTTGCTCTCTGTGGATTTCCAGAAAGTGTACTGTGG	787				
Db	725	AGAAACAGAAATTCAGAGGACTTGCTCTCTGTGGATTTCCAGAAAGTGTACTGTGG	784				
Qy	788	TCTCTTTTCAAGACAGAAATACCTTAGTCTCAGCAGAGCTGTGGATGGGATTAATCAAG	847				
Db	785	TCTCTTTTCAAGACAGAAATACCTTAGTCTCAGCAGAGCTGTGGATGGGATTAATCAAG	844				
Qy	848	TATGGGATTTAGTAAGATTTATCTGTCTTATCGACAGAACCCATAGCATCCAAGTCTT	907				
Db	845	TATGGGATTTAGTAAGATTTATCTGTCTTATCGACAGAACCCATAGCATCCAAGTCTT	904				
Qy	908	TCTGTGACCCAGGTAGCAGACTCGAAACTTGGATTTCAAGTCTGATTTTGGGATTTCA	967				
Db	905	TCTGTGACCCAGGTAGCAGACTCGAAACTTGGATTTCAAGTCTGATTTTGGGATTTCA	964				
Qy	968	CTGGCTCTACTTTTATTTGCTAATTGCAACAGGATAACATCTACATGTTTAAATGACTG	1027				
Db	965	CTGGCTCTACTTTTATTTGCTAATTGCAACAGGATAACATCTACATGTTTAAATGACTG	1024				

Qy	1028	GTTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACCAAGAACTCTACCTTTTATGTAA	1087				
Db	1025	GTTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACCAAGAACTCTACCTTTTATGTAA	1084				
Qy	1088	AATCCAGCCTTAGTCCAGATGACCAAGTTTATTTAGTCAGTGGCTCAAGTGATGAAGCTGCCT	1147				
Db	1085	AATCCAGCCTTAGTCCAGATGACCAAGTTTATTTAGTCAGTGGCTCAAGTGATGAAGCTGCCT	1144				
Qy	1148	ACATATGGAAGGTCTCCACACCCCTGGCAACCTCTCTACTGTGCTCTCTGGGTCAATCTCAAG	1207				
Db	1145	ACATATGGAAGGTCTCCACACCCCTGGCAACCTCTCTACTGTGCTCTCTGGGTCAATCTCAAG	1204				
Qy	1208	AGGTCAAGTCTGTGCTGTGGTGTCTATCTGACTTTCACAAAGATTTGCTTACCTGTCTGATG	1267				
Db	1205	AGGTCAAGTCTGTGCTGTGGTGTCTATCTGACTTTCACAAAGATTTGCTTACCTGTCTGATG	1264				
Qy	1268	ACAATACACATAAATACTGGCGCTTGAATAGAGGCTTTAGAGGAGAAACCCAGGAGGTGATA	1327				
Db	1265	ACAATACACATAAATACTGGCGCTTGAATAGAGGCTTTAGAGGAGAAACCCAGGAGGTGATA	1324				
Qy	1328	AACCTTCCACGCTGGTGGGCTCTCAGAAAGAAAGAGTCAAGACCTGGCCCTAGTAA	1387				
Db	1325	AACCTTCCACGCTGGTGGGCTCTCAGAAAGAAAGAGTCAAGACCTGGCCCTAGTAA	1384				
Qy	1388	CAGTAAAGAGTACAGAGTACTCTGCAAGAGCCCCAGGGTAAAGTGCAATCCATCCA	1447				
Db	1385	CAGTAAAGAGTACAGAGTACTCTGCAAGAGCCCCAGGGTAAAGTGCAATCCATCCA	1444				
Qy	1448	ATTCTTCCCGCTCATCCGCGACTTGTGCCCAAGCTGTGTGGAGACCTCTCTCTCTCT	1507				
Db	1445	ATTCTTCCCGCTCATCCGCGACTTGTGCCCAAGCTGTGTGGAGACCTCTCTCTCTCTCT	1504				
Qy	1508	CAAAATCTCTACGTTCTCTATTTAAACCTCTCTCTGCAAGGCCCGCTCTCCATCAACA	1567				
Db	1505	CAAAATCTCTACGTTCTCTATTTAAACCTCTCTCTGCAAGGCCCGCTCTCCATCAACA	1564				
Qy	1568	GAAGAGCTCTGTCTCTCGCTCTCTCCAGGACCACTTCACTTTTCAAGATGTGATTA	1627				
Db	1565	GAAGAGCTCTGTCTCTCGCTCTCTCCAGGACCACTTCACTTTTCAAGATGTGATTA	1624				
Qy	1628	GAAGCTGGGTGACCGCAACCTTCTCTCATCAACACCCATCACTCCACCTCTCTCGGAGA	1687				
Db	1625	GAAGCTGGGTGACCGCAACCTTCTCTCATCAACACCCATCACTCCACCTCTCTCGGAGA	1684				
Qy	1688	CCAAGATCATGTCTCCGAGAAAAGCCCTTATTTCTGTGAGCCAGAGTCATCCCAAGCAG	1747				
Db	1685	CCAAGATCATGTCTCCGAGAAAAGCCCTTATTTCTGTGAGCCAGAGTCATCCCAAGCAG	1744				
Qy	1748	AGGCTTGTCTGAGTCTAGAAATAGAGTAAAGAGGAGCTAGACTCAAGCTGTCTGGAGA	1807				
Db	1745	AGGCTTGTCTGAGTCTAGAAATAGAGTAAAGAGGAGCTAGACTCAAGCTGTCTGGAGA	1804				
Qy	1808	GTGTGAAAACAAAAGTGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGATGCCAAGTTG	1867				
Db	1805	GTGTGAAAACAAAAGTGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGATGCCAAGTTG	1864				
Qy	1868	AAAATCTTTCATTTGGATCTGTGCTTCCCTTGTGTAACCAAGAGACCTTTAGTAGAGACT	1927				
Db	1865	AAAATCTTTCATTTGGATCTGTGCTTCCCTTGTGTAACCAAGAGACCTTTAGTAGAGACT	1924				
Qy	1928	CTCTAGGTCCTACCAATCAAGCAAAATTTGAAGGAGCTGGTACCAGATATCTCAGAGCCCTC	1987				
Db	1925	CTCTAGGTCCTACCAATCAAGCAAAATTTGAAGGAGCTGGTACCAGATATCTCAGAGCCCTC	1984				
Qy	1988	CGTCTCTATCAGTCCGCTATGCTTTCAGAAAGCTGTGGAAACGCTACCTCTTCTTTGAGAC	2047				
Db	1985	CGTCTCTATCAGTCCGCTATGCTTTCAGAAAGCTGTGGAAACGCTACCTCTTCTTTGAGAC	2044				
Qy	2048	CTTGTGGAGAGGGTCTGAAATGGTAGGCAAGAGATAGTTTCCCAGAGATATAAAACT	2107				
Db	2045	CTTGTGGAGAGGGTCTGAAATGGTAGGCAAGAGATAGTTTCCCAGAGATATAAAACT	2104				
Qy	2108	GGTTGTTGGCCATGGCAGCCCAACCGAAGGCTGAGATCCATCTCCACGAAGTCCGTCTCAT	2167				

Qy	1023	GACTGGGTTGAAGACTTTCTCCAGTGGCTATTTTCAATGGAACACAGAACTCTACCTTTTA	1082
Db	1056	GACTGGGTTGAAGACTTTCTCCAGTGGCTATTTTCAATGGAACACAGAACTCTACCTTTTA	1115
Qy	1083	TGTAANAATCCAGCCCTTAGTCCAGATACCAAGTTTTTTTAGTCAGTGGCTCAAGTATGAAGC	1142
Db	1116	TGTAANAATCCAGCCCTTAGTCCAGATACCAAGTTTTTTTAGTCAGTGGCTCAAGTATGAAGC	1175
Qy	1143	TGCCTACATATGGAAGGCTCCACACCCCTGGCAACCTCCTACTGTGTCTCCTGGGTCAATTC	1202
Db	1176	TGCCTACATATGGAAGGCTCCACACCCCTGGCAACCTCCTACTGTGTCTCCTGGGTCAATTC	1235
Qy	1203	TCAAGAGGTCAAGCTGTGTGTGTGGTGTCCATCTGAATTCAAAAAGATGCTACCTGTGTTTC	1262
Db	1236	TCAAGAGGTCAAGCTGTGTGTGTGGTGTCCATCTGAATTCAAAAAGATGCTACCTGTGTTTC	1295
Qy	1263	TGATGCAATACACTAAANAATCTGGCCCTTGAATAGAGGCTTAGAGAGAGAAAACGAGGAGG	1322
Db	1296	TGATGCAATACACTAAANAATCTGGCCCTTGAATAGAGGCTTAGAGAGAGAAAACGAGGAGG	1355
Qy	1323	TGATAAACTTTTCCACCGGTGGGTTGGGCTCTCTCAGAAAGAAAAGAGTCAAGACCTGGCCCT	1382
Db	1356	TGATAAACTTTTCCACCGGTGGGTTGGGCTCTCTCAGAAAGAAAAGAGTCAAGACCTGGCCCT	1415
Qy	1383	AGTAACAGTAAACGATAGCCAGAGTACTCTGTGCCAAAAGCCCCAGGGTAAAGTGCATATCC	1442
Db	1416	AGTAACAGTAAACGATAGCCAGAGTACTCTGTGCCAAAAGCCCCAGGGTAAAGTGCATATCC	1475
Qy	1443	ATCCAAATCTTCCCGCTCATCCGAGCTGTGCCCCAAGCTGTGCTGGAGACCTCCCTCT	1502
Db	1476	ATCCAAATCTTCCCGCTCATCCGAGCTGTGCCCCAAGCTGTGCTGGAGACCTCCCTCT	1535
Qy	1503	TCCTTTCAAAATCTCCTACGTTCTCTATTAAAACTCTCTGCCAAAGCCCGGCTCTCCCAT	1562
Db	1536	TCCTTTCAAAATCTCCTACGTTCTCTATTAAAACTCTCTGCCAAAGCCCGGCTCTCCCAT	1595
Qy	1563	CAACAGAAAGGCTCTGTCTCCTCCGTCTCTCCCAAGCCACTTCATCTTTCAAGATGTC	1622
Db	1596	CAACAGAAAGGCTCTGTCTCCTCCGTCTCTCCCAAGCCACTTCATCTTTCAAGATGTC	1655
Qy	1623	GATTAGAAACTGGGTGACCCGAAACACTTCTCTCATCACCCCACTCACTCCACTGCTTC	1682
Db	1656	GATTAGAAACTGGGTGACCCGAAACACTTCTCTCATCACCCCACTCACTCCACTGCTTC	1715
Qy	1683	GGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTCTGTGAGCCAGAAAGTCAATCCCA	1742
Db	1716	GGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTCTGTGAGCCAGAAAGTCAATCCCA	1775
Qy	1743	AGCAGAGGCTTGCTCTGAGTCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCT	1802
Db	1776	AGCAGAGGCTTGCTCTGAGTCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCT	1835
Qy	1803	GGAGAGTGTGAAAACAAAAGCTGTGTGAAGAGTGTGTAACTGTGTGACTCAGCTTGTATGGCCA	1862
Db	1836	GGAGAGTGTGAAAACAAAAGCTGTGTGAAGAGTGTGTAACTGTGTGACTCAGCTTGTATGGCCA	1895
Qy	1863	AGTTGAAAATCTTCATTTTGGATCTGTGCTGCCCTTGTGGTAAACAGAGAAAGCTTTAGTAA	1922
Db	1896	AGTTGAAAATCTTCATTTTGGATCTGTGCTGCCCTTGTGGTAAACAGAGAAAGCTTTAGTAA	1955
Qy	1923	GGACTCTTAGGTCCTACCAATCAAGCAAAATTTGAGGAGGCTGGTACCAAGTATCTCAGA	1982
Db	1956	GGACTCTTAGGTCCTACCAATCAAGCAAAATTTGAGGAGGCTGGTACCAAGTATCTCAGA	2015
Qy	1983	GCCTCCGCTCTCCTATCAGTCCGATGCTTTCAGAAAAGCTGTGTGAAAGCTCACTCTTCCTTT	2042
Db	2016	GCCTCCGCTCTCCTATCAGTCCGATGCTTTCAGAAAAGCTGTGTGAAAGCTCACTCTTCCTTT	2075
Qy	2043	GAGACCTTTGTGGAGAGGGTCTGAAATGGTATAGGCAAAAGAGAAATAGTTCCTCCAGAAATAA	2102
Db	2076	GAGACCTTTGTGGAGAGGGTCTGAAATGGTATAGGCAAAAGAGAAATAGTTCCTCCAGAAATAA	2135

Qy	2103	AAACTGGTTGTTGGCCATGCGAGCCTAAACGGAAGGCTGAGAATCCATCTCTCCACGAAGTCC	21616
Db	2136	AAACTGGTTGTTGGCCATGCGAGCCTAAACGGAAGGCTGAGAATCCATCTCTCCACGAAGTCC	21955
Qy	2163	GTCAATCCAGACACCCTAATTTCAGAGGAGACAGAGCGGAAGACATGTCGAAGCCCGGTTCAC	2222
Db	2196	GTCAATCCAGACACCCTAATTTCAGAGGAGACAGAGCGGAAGACATGTCGAAGCCCGGTTCAC	22555
Qy	2223	CATCACGCCCAGCTCCATGAGGAAAATCTGCACATACCTCCATAGAAAAGTCCCAGGAGGA	22822
Db	2256	CATCACGCCCAGCTCCATGAGGAAAATCTGCACATACCTCCATAGAAAAGTCCCAGGAGGA	2315
Qy	2283	CTTCTGTGGTCTCGTAACACACTCAACNGAATATATAGATTCTTAATCTGAGTGAGTTACTCGGC	2342
Db	2316	CTTCTGTGGTCTCGTAACACACTCAACNGAATATATAGATTCTTAATCTGAGTGAGTTACTCGGC	2375
Qy	2343	TTTGGTCCACTAAAAACAAGCTGAGCTTTGGTCCACTAAAAACAAGTCAAATAACAAGAG	2402
Db	2376	TTTGGTCCACTAAAAACAAGCTGAGCTTTGGTCCACTAAAAACAAGTCAAATAACAAGAG	2435
Qy	2403	TGACTCTATAACTCTGTGCTTTTAGAAAGCTGGCTTTTTTCATTTTTTAGACAAAATCTTTTC	2462
Db	2436	TGACTCTATAACTCTGTGCTTTTAGAAAGCTGGCTTTTTTCATTTTTTAGACAAAATCTTTTC	2495
Qy	2463	AACGCTGAATGTACCTAATCTGGTCTCTACTACCATATGCTATATGTCAGCTTCCCGAGGA	2522
Db	2496	AACGCTGAATGTACCTAATCTGGTCTCTACTACCATATGCTATATGTCAGCTTCCCGAGGA	2555
Qy	2523	TGAATGCTGTGTTTAAATTTTCAANAAGTAAATTTGTCACTAGCATTTTGAATGAATAG	2582
Db	2556	TGAATGCTGTGTTTAAATTTTCAANAAGTAAATTTGTCACTAGCATTTTGAATGAATAG	2615
Qy	2583	TCTTCACTTTTTTAAATTAATTCATCTCTCTATATAATATGATATGATATGAGGAGCA	2642
Db	2616	TCTTCACTTTTTTAAATTAATTCATCTCTCTATATAATATGATATGATATGAGGAGCA	2675
Qy	2643	AAAACAAGTTTCTGTGTTATCCCTGAAACTTTCTATGCTCAGTGTGAAAGTATCTCCAGCC	2702
Db	2676	AAAACAAGTTTCTGTGTTATCCCTGAAACTTTCTATGCTCAGTGTGAAAGTATCTCCAGCC	2735
Qy	2703	ACAGCATGAGGCTGTGAAGGCTGATGTGAGAAATCTCTGCTGAAGACCCCTGGTTCTGT	2762
Db	2736	ACAGCATGAGGCTGTGAAGGCTGATGTGAGAAATCTCTGCTGAAGACCCCTGGTTCTGT	2795
Qy	2763	TCTGCCTCCAACATGTATAATTTTATTTGAAATACATAATCTTTTCACTA	2812
Db	2796	TCTGCCTCCAACATGTATAATTTTATTTGAAATACATAATCTTTTCACTA	2845
RESULT	9		
AX926533			
LOCUS	AX926533	4221 bp	DNA linear PAT 19-DEC-2003
DEFINITION	Sequence 85 from Patent EP1355150.		
ACCESSION	AX926533		
VERSION	AX926533.1	GI:40246344	
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominoidea; Homo.		
REFERENCE	1.	4221	
AUTHORS	Wang, Y.		
TITLE	Panel of nucleic acid sequences for cancer diagnosis		
JOURNAL	Patent: EP 1355150-A 85 22-OCT-2003;		
FEATURES	Ortho-Clinical Diagnostics, Inc. (US) Location/Qualifiers 1..4221 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"		
source			
ORIGIN			

Query Match	98.5%	Score 2789.8	DB 2	Length 4221	
Best Local Similarity	99.9%	Pred. No. 0			
Matches 2791	Conservative 0	Mismatches 2	Indels 0	Gaps 0	
Qy	22	CGATAACGATTTTGTGTGTGAGAGGCGCAACGTCGCGATTTCTGCTGAACATTGGAGGCGATT	81		
Db	1	CGATAACGATTTTGTGTGTGAGAGGCGCAACGTCGCGATTTCTGCTGAACATTGGAGGCGATT	60		
Qy	82	TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG	141		
Db	61	TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG	120		
Qy	142	CTCCGCAGACCCAGCTTGGGCTCTGAGAAATGGATGGTCTTCAAAATACCCTCTTCAA	201		
Db	121	CTCCGCAGACCCAGCTTGGGCTCTGAGAAATGGATGGTCTTCAAAATACCCTCTTCAA	180		
Qy	202	TCCCTTCTGACTGGTATCAGTCAGTGGTAAATGATGAACACACTTCTTATGGAGAAACA	261		
Db	181	TCCCTTCTGACTGGTATCAGTCAGTGGTAAATGATGAACACACTTCTTATGGAGAAACA	240		
Qy	262	GGAGTCCCAGTTCTCTTTTGGATGACCTTCTTCTTCTGCTCCCAATATGGAAACATGTA	321		
Db	241	GGAGTCCCAGTTCTCTTTTGGATGACCTTCTTCTTCTGCTCCCAATATGGAAACATGTA	300		
Qy	322	CTACGAGTGGCAATGAAGAAGCTTTGTTCGATTGTATAACACAGAAATCACAAAGTTTC	381		
Db	301	CTACGAGTGGCAATGAAGAAGCTTTGTTCGATTGTATAACACAGAAATCACAAAGTTTC	360		
Qy	382	AGAAAGAGTGCTTCAAGAATGATGGCTCACGTGGAATGCCGTCCTTTGACCTGGCCCTGG	441		
Db	361	AGAAAGAGTGCTTCAAGAATGATGGCTCACGTGGAATGCCGTCCTTTGACCTGGCCCTGG	420		
Qy	442	GTTCCTGGTGAACTTAAACTTTGTACAGCAGCAGGTGATCAAAACAGCCAAATTTTGGGAC	501		
Db	421	GTTCCTGGTGAACTTAAACTTTGTACAGCAGCAGGTGATCAAAACAGCCAAATTTTGGGAC	480		
Qy	502	GTAAGAGCTGGTGAGCTGATTTGGAACATGCAAGGTCACTAATGCAGCCTCAAGTCAGTT	561		
Db	481	GTAAGAGCTGGTGAGCTGATTTGGAACATGCAAGGTCACTAATGCAGCCTCAAGTCAGTT	540		
Qy	562	GCTTTTCTAAGTTTGAAAGCTGATTTCTGTACGGGTGGAAGAGATGCAACATTATG	621		
Db	541	GCTTTTCTAAGTTTGAAAGCTGATTTCTGTACGGGTGGAAGAGATGCAACATTATG	600		
Qy	622	GTCTGGGATACAGGTGCAACAAAAAGATGGGTTTTATAGGCAAGTGAATCAAAATCAGT	681		
Db	601	GTCTGGGATACAGGTGCAACAAAAAGATGGGTTTTATAGGCAAGTGAATCAAAATCAGT	660		
Qy	682	GGAGCTCAAAATCCTCAGACAGCAACCCCTTCAAAACCCAAAGAGAAACAGAAATCA	741		
Db	661	GGAGCTCAAAATCCTCAGACAGCAACCCCTTCAAAACCCAAAGAGAAACAGAAATCA	720		
Qy	742	AAAGGACTTGCTCTCTGTGGATTTCCAGCAAGTGTACTGTGGTCTCTTTCAAGAC	801		
Db	721	AAAGGACTTGCTCTCTGTGGATTTCCAGCAAGTGTACTGTGGTCTCTTTCAAGAC	780		
Qy	802	GAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAAGTATGGGATTTACGT	861		
Db	781	GAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAAGTATGGGATTTACGT	840		
Qy	862	AAGAAATATATCTGCTTATCGACAAGAACCCATAGCATCCAAAGTCTTCTGTATCCCAAGT	921		
Db	841	AAGAAATATATCTGCTTATCGACAAGAACCCATAGCATCCAAAGTCTTCTGTATCCCAAGT	900		
Qy	922	AGCAGCACTCGAAAACCTTGGATTTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA	981		
Db	901	AGCAGCACTCGAAAACCTTGGATTTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA	960		
Qy	982	TTTGCTAATTGCACAGACGATAACATCTACATGTTTAAATATGACTGGGTGAAAGACTTCT	1041		
Db	961	TTTGCTAATTGCACAGACGATAACATCTACATGTTTAAATATGACTGGGTGAAAGACTTCT	1020		
Qy	1042	CCAGTGGCTATTTTCAATGGCACACCGAACTCTACCTTTTATGTAAATACGCTTAGT	1101		

Db	2101	GCAGCCAAACGGAAGCTGAGAAATCCATCTCCAGAAAGTCCGTCATCCAGACACCCCAAT	2160
Qy	2182	TCCAGGACAGAGCGGAAAGACATTTGCCAAGCCCGGTCAACATCAACGCCAGCTCCATG	2241
Db	2161	TCCAGGACAGAGCGGAAAGACATTTGCCAAGCCCGGTCAACATCAACGCCAGCTCCATG	2220
Qy	2242	AGGAAATCTGCACATACTTCCATAGAAAGTCCAGGAGGACTTCTGTGGTCTCTGAAAC	2301
Db	2221	AGGAAATCTGCACATACTTCCATAGAAAGTCCAGGAGGACTTCTGTGGTCTCTGAAAC	2280
Qy	2302	TCAACAGAAATATAGATCTTAATCTGAGTGAGTACTGAGCTTTTGGTCCACCTAATAACAAG	2361
Db	2281	TCAACAGAAATATAGATCTTAATCTGAGTGAGTACTGAGCTTTTGGTCCACCTAATAACAAG	2340
Qy	2362	CTGAGCTTTTGGTCCACTAAACCAAGATGAAATAACAAGAGTGACTCTATAACTCTGGTC	2421
Db	2341	CTGAGCTTTTGGTCCACTAAACCAAGATGAAATAACAAGAGTGACTCTATAACTCTGGTC	2400
Qy	2422	TTTAAGAAAGCTGCCCTTTTTCATTTTATAGACAAATCTTTTCAACGCTGAATGTACCTAA	2481
Db	2401	TTTAAGAAAGCTGCCCTTTTTCATTTTATAGACAAATCTTTTCAACGCTGAATGTACCTAA	2460
Qy	2482	TCTGGTCTTACTACCATATATGATATGACGTTTCCGAGGATGAATGCTGTGTTTAAATT	2541
Db	2461	TCTGGTCTTACTACCATATATGATATGACGTTTCCGAGGATGAATGCTGTGTTTAAATT	2520
Qy	2542	TCATAAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATPAT	2601
Db	2521	TCATAAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATPAT	2580
Qy	2602	TCATCTTCTATATATATGACATCCAGTTCATCGAGGCAAAAACAAGTTCTTGTTA	2661
Db	2581	TCATCTTCTATATATATGACATCCAGTTCATCGAGGCAAAAACAAGTTCTTGTTA	2640
Qy	2662	TCCTGAACTTCTATGCTCAGTGGAAGATATCTGCCAGCCACAGCATGAGGCCGTGAA	2721
Db	2641	TCCTGAACTTCTATGCTCAGTGGAAGATATCTGCCAGCCACAGCATGAGGCCGTGAA	2700
Qy	2722	GGCTGACTGAGAAATCTCTGCTGTAAGACCCCTGGTCTGTTCTGCCCTCCAACATGTATA	2781
Db	2701	GGCTGACTGAGAAATCTCTGCTGTAAGACCCCTGGTCTGTTCTGCCCTCCAACATGTATA	2760
Qy	2782	ATTTTATTTGAAATACATAATCTTTTCACTATG	2814
Db	2761	ATTTTATTTGAAATACATAATCTTTTCACTATG	2793
RESULT 10			
AX951786			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Homo sapiens (human)			
Homo sapiens			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
Hominidae; Homo.			
1			
Wang, Y.			
Markers for breast cancer prognosis			
TITLE			
Patent: EP 1367138-A 85 03-DEC-2003;			
JOURNAL			
Ortho Clinical Diagnostics Inc. (US)			
FEATURES			
Location/Qualifiers			
1..4221			
source			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
ORIGIN			
Query Match			
, Best Local Similarity			
98.5%; Score 2789.8; DB 2; Length 4221;			
99.9%; Pred. No. 0;			

Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	22	CGATAACGATTTGTGTGTGAGAGCGCAAGCTGCGATTTCTGCTGAACCTTGGAGGCAATT	81
Db	1	CGATAACGATTTGTGTGTGAGAGCGCAAGCTGCGATTTCTGCTGAACCTTGGAGGCAATT	60
Qy	82	TCTACGACTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG	141
Db	61	TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG	120
Qy	142	CTCCGCCAGCCAGCTTTGGCGTCTCGAGAAATGGATGGTCTTTCACAAATACCTCTTCAA	201
Db	121	CTCCGCCAGCCAGCTTTGGCGTCTCGAGAAATGGATGGTCTTTCACAAATACCTCTTCAA	180
Qy	202	TCCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAAACACATTTCTTATGGAACA	261
Db	181	TCCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAAACACATTTCTTATGGAACA	240
Qy	262	GGAGTCCCAGTTCCTCCTTTTGGATGTACCTTCTCTTCTGCTCCCAATATGGAACATGTA	321
Db	241	GGAGTCCCAGTTCCTCCTTTTGGATGTACCTTCTCTTCTGCTCCCAATATGGAACATGTA	300
Qy	322	CTAGCAGTTGCCAATGAAGAGGCTTTGTTCGATTTGTAACACAGAAATCAAAAGTTTC	381
Db	301	CTAGCAGTTGCCAATGAAGAGGCTTTGTTCGATTTGTAACACAGAAATCAAAAGTTTC	360
Qy	382	AGAAAGAGTGGCTTCAAAGAAATGGATGGCTCACCTGGAATGCCGTCTTTGACCTGGCTGG	441
Db	361	AGAAAGAGTGGCTTCAAAGAAATGGATGGCTCACCTGGAATGCCGTCTTTGACCTGGCTGG	420
Qy	442	GTTCTCGTGGAACCTTAAACCTTTTACAGCAGCAGTGATCAAAAGAGCAAAATTTTGGGAC	501
Db	421	GTTCTCGTGGAACCTTAAACCTTTTACAGCAGCAGTGATCAAAAGAGCAAAATTTTGGGAC	480
Qy	502	GTAAGAGCTGGTGAAGCTGATTTGGAACATGCAAAAGGTCAATCAATGCAGCCTCAAGTCAGTT	561
Db	481	GTAAGAGCTGGTGAAGCTGATTTGGAACATGCAAAAGGTCAATCAATGCAGCCTCAAGTCAGTT	540
Qy	562	GCCTTTTCTAAGTTTGAAGAAAGCTGTATTTCTGTACGGGTGGAAGAGATGGCAACATTATG	621
Db	541	GCCTTTTCTAAGTTTGAAGAAAGCTGTATTTCTGTACGGGTGGAAGAGATGGCAACATTATG	600
Qy	622	GTCTGGGATACAGGTGCACAAAGATGGGTTTTATAGGCAAGTGAATCAATCAATCAGT	681
Db	601	GTCTGGGATACAGGTGCACAAAGATGGGTTTTATAGGCAAGTGAATCAATCAATCAGT	660
Qy	682	GGAGCTCAATATACCTCAGCAACAAACCCCTTTCAAAACCCCAAGAGAAACAGAAATTC	741
Db	661	GGAGCTCAATATACCTCAGCAACAAACCCCTTTCAAAACCCCAAGAGAAACAGAAATTC	720
Qy	742	AAAGGACTTGTCTCTTCTGTGGATTTCCAGCAAAAGTTTACTGTGGTCTCTTTCAAGAC	801
Db	721	AAAGGACTTGTCTCTTCTGTGGATTTCCAGCAAAAGTTTACTGTGGTCTCTTTCAAGAC	780
Qy	802	GAGAAATACCTTAGTCTCAGCAGGAGCTGTGATGGGTAATCAAGTATGGGATTTAGT	861
Db	781	GAGAAATACCTTAGTCTCAGCAGGAGCTGTGATGGGTAATCAAGTATGGGATTTAGT	840
Qy	862	AAGAAATATACCTGTCTATCGACAAAGAACCCATAGCATCCAAAGTCTTTCTCTACCCAGGT	921
Db	841	AAGAAATATACCTGTCTATCGACAAAGAACCCATAGCATCCAAAGTCTTTCTCTACCCAGGT	900
Qy	922	AGCAGCCTCGAAAACTTTGGATATTCAGTCTGATTTTGGATTTCCACTGGCTCTACTTTTA	981
Db	901	AGCAGCCTCGAAAACTTTGGATATTCAGTCTGATTTTGGATTTCCACTGGCTCTACTTTTA	960
Qy	982	TTTCTGTAATTCGACAGAGCAATCATCTACATGTTTAAATATGATGGTTCGAAGACTTCT	1041
Db	961	TTTCTGTAATTCGACAGAGCAATCATCTACATGTTTAAATATGATGGTTCGAAGACTTCT	1020
Qy	1042	CCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTATGTAATAATCCAGCCTTAGT	1101
Db	1021	CCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTATGTAATAATCCAGCCTTAGT	1080

QY 22 CGATAACGATTTGTTGTGAGAGCGGCAACGTGCGATTTCTGCTGAATCTTGAGGCATT 81
Db 1 CGATAACGATTTGTTGTGAGAGCGCAAGCTGCGATTTCTGCTGAATCTTGAGGCATT 60
QY 82 TCTAGACATTTTCTCTCAGCTGAGGCTTTTCTCCGAGCCCTGATGCTCTTCAATTCGGTG 141
Db 61 TCTAGACATTTTCTCTCAGCTGAGGCTTTTCTCCGAGCCCTGATGCTCTTCAATTCGGTG 120
QY 142 TTCGCCACGCCAGCTTGGCGTCTGAGAAATGGATGGTCTTCAATATCCCTCTTCAA 201
Db 121 CTCGCCACGCCAGCTTGGCGTCTGAGAAATGGATGGTCTTCAATATCCCTCTTCAA 180
QY 202 TCCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACACACTTCTTATGGGAACA 261
Db 181 TCCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACACACTTCTTATGGGAACA 240
QY 262 GGAGTCCCAAGTTCCCTCTTTGGATGTACCTTCTCTCTGCTCCCAATATGGAAATGTA 321
Db 241 GGAGTCCCAAGTTCCCTCTTTGGATGTACCTTCTCTCTGCTCCCAATATGGAAATGTA 300
QY 322 CTAGCAGTTGCCAATGAAGAAGGCTTTGTTCGATTTGATATAACACAGAATCAAAAGTTTC 381
Db 301 CTAGCAGTTGCCAATGAAGAAGGCTTTGTTCGATTTGATATAACACAGAATCAAAAGTTTC 360
QY 382 AGAAAGAGTGTCTCAAGAAATGGATGGCTCACTGGAATGCGGTCTTGAACCTGCCCTGG 441
Db 361 AGAAAGAGTGTCTCAAGAAATGGATGGCTCACTGGAATGCGGTCTTGAACCTGCCCTGG 420
QY 442 GTTCTGTTGAACTTTAAATCTGTACAGCAGCAGGTGATCAACAGCCAAATTTTGGGAC 501
Db 421 GTTCTGTTGAACTTTAAATCTGTACAGCAGCAGGTGATCAACAGCCAAATTTTGGGAC 480
QY 502 GTAAAAGCTGGTGAGCTGATTTGGAACATGCCAAGGTCAATGAGCCTCAAGTCACTT 561
Db 481 GTAAAAGCTGGTGAGCTGATTTGGAACATGCCAAGGTCAATGAGCCTCAAGTCACTT 540
QY 562 GCCTTTCTAAGTTTGAGAAAGCTGATTTCTGTACGGGTGGAGAGATGGCAACATTATG 621
Db 541 GCCTTTCTAAGTTTGAGAAAGCTGATTTCTGTACGGGTGGAGAGATGGCAACATTATG 600
QY 622 GTCTGGGATACCAAGTGCACAAAAAGATGGTTTTATAGGCAAGTGAATCAAAATCAGT 681
Db 601 GTCTGGGATACCAAGTGCACAAAAAGATGGTTTTATAGGCAAGTGAATCAAAATCAGT 660
QY 682 GGAGCTCACAATPACTCAGACAAGCAAAACCCCTTCAAAAACCCCAAGAGAAAAGAAATCA 741
Db 661 GGAGCTCACAATPACTCAGACAAGCAAAACCCCTTCAAAAACCCCAAGAGAAAAGAAATCA 720
QY 742 AAAGGACTTGCTCTCTGTGGATTTCCAGCAAGTGTACTGTGGTCTCTTTCAGAC 801
Db 721 AAAGGACTTGCTCTCTGTGGATTTCCAGCAAGTGTACTGTGGTCTCTTTCAGAC 780
QY 802 GAGAAATACCTTAGTCTCAGCAGAGCTGTGGATGGGTAATCAAAAGTATGGGATTTACGT 861
Db 781 GAGAAATACCTTAGTCTCAGCAGAGCTGTGGATGGGTAATCAAAAGTATGGGATTTACGT 840
QY 862 AAGAAATTAATPACTGTTATCGACAAGAACCCATPAGCATCCAAGTCTTTCTGTACCCAGGT 921
Db 841 AAGAAATTAATPACTGTTATCGACAAGAACCCATPAGCATCCAAGTCTTTCTGTACCCAGGT 900
QY 922 AGCAGACTCGAAATCTGGATTAATCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 981
Db 901 AGCAGACTCGAAATCTGGATTAATCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 960
QY 982 TTTGCTTAATGACAGACGATAAATCTACATGTTTAAATATACATGCTGGGTGGAAGCTTCT 1041
Db 961 TTTGCTTAATGACAGACGATAAATCTACATGTTTAAATATATGACATGCTGGGTGGAAGCTTCT 1020
QY 1042 CCAGTGGCTATTTTCAATGGACACCAAGAACTCTACCTTTTATGTAAAAATCCAGCCTTAGT 1101
Db 1021 CCAGTGGCTATTTTCAATGGACACCAAGAACTCTACCTTTTATGTAAAAATCCAGCCTTAGT 1080

QY 1102 CCAGATGACAGATTTTGTAGTCAGTGGCTCAAGTGAAGCTGCTACATATGAAGGTC 1161
Db 1081 CCAGATGACAGATTTTGTAGTCAGTGGCTCAAGTGAAGCTGCTACATATGAAGGTC 1140
QY 1162 TCCACACCCCTGGCAACCTCTCTACTGTCTCTCTGGGTCAATTTCTCAAGAGGTCACTCTGTG 1221
Db 1141 TCCACACCCCTGGCAACCTCTCTACTGTCTCTCTGGGTCAATTTCTCAAGAGGTCACTCTGTG 1200
QY 1222 TGCTGGTGTGCATCTGATCTTCAAAAGATTTGCTACTCTGTCTGATGACAAATACATAAAA 1281
Db 1201 TGCTGGTGTGCATCTGATCTTCAAAAGATTTGCTACTCTGTCTGATGACAAATACATAAAA 1260
QY 1282 ATCTGGCGCTTGAATAGAGGCTTAGAGAGAAACAGAGAGGTATAAATTTTCAACGGTG 1341
Db 1261 ATCTGGCGCTTGAATAGAGGCTTAGAGAGAAACAGAGAGGTATAAATTTTCAACGGTG 1320
QY 1342 GGTTCGGCTCTTCAGAAAGAAAAAGAGTCAAGACCTGGCCTAGTAAACAGTAAACGAGTAGC 1401
Db 1321 GGTTCGGCTCTTCAGAAAGAAAAAGAGTCAAGACCTGGCCTAGTAAACAGTAAACGAGTAGC 1380
QY 1402 CAGAGTACTCTGCAAAAGCCCCCAGGGTAAAGTGCAATCCATCCAATTTCTTCCCGCTCA 1461
Db 1381 CAGAGTACTCTGCAAAAGCCCCCAGGGTAAAGTGCAATCCATCCAATTTCTTCCCGCTCA 1440
QY 1462 TCCGACAGCTGTGCCCCCAAGCTGTCTGGAGACCTCCCTCTTCAAATFACTCTCTACG 1521
Db 1441 TCCGACAGCTGTGCCCCCAAGCTGTCTGGAGACCTCCCTCTTCAAATFACTCTCTACG 1500
QY 1522 TTCTCTATTAATAACCTCTCTGCCNAGGCCCGGTCTCCCATCAACAGAGAGGCTCTGTCT 1581
Db 1501 TTCTCTATTAATAACCTCTCTGCCNAGGCCCGGTCTCCCATCAACAGAGAGGCTCTGTCT 1560
QY 1582 TCCTCCGCTCTCTCCCAAGCCACCTTCATCTTTCAAGATGTGATTAGAACTGGGTGACC 1641
Db 1561 TCCTCCGCTCTCTCCCAAGCCACCTTCATCTTTCAAGATGTGATTAGAACTGGGTGACC 1620
QY 1642 GGAACACTTCTCATCAACACCCNCTCACTCCACTGCTTCGGAGACCAAGATCATGTCT 1701
Db 1621 GGAACACTTCTCATCAACACCCNCTCACTCCACTGCTTCGGAGACCAAGATCATGTCT 1680
QY 1702 CCGAGAAAAAGCCCTTATTTCTGTGAGCCAGAAGTCATCCCAAGCAGAGGCTTGTCTGAG 1761
Db 1681 CCGAGAAAAAGCCCTTATTTCTGTGAGCCAGAAGTCATCCCAAGCAGAGGCTTGTCTGAG 1740
QY 1762 TCTAGAAATAGAGTAAAGAGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAG 1821
Db 1741 TCTAGAAATAGAGTAAAGAGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAG 1800
QY 1822 TGTGTGAAGAGTGTCTGAGCTGAGCTTCATGGCCCAAGTTGAAAAATCTTCATTTG 1881
Db 1801 TGTGTGAAGAGTGTCTGAGCTGAGCTTCATGGCCCAAGTTGAAAAATCTTCATTTG 1860
QY 1882 GATCTGTCTGCTTGTCTGAGTAAACAGAGAACCTTATAGTAAGGACTCTCTAGGCTCTTACC 1941
Db 1861 GATCTGTCTGCTTGTCTGAGTAAACAGAGAACCTTATAGTAAGGACTCTCTAGGCTCTTACC 1920
QY 1942 AAATCAAGCAAAATTAAGAGAGCTGTACAGTATCTCAGAGCCTCCGCTCTCTATCAGT 2001
Db 1921 AAATCAAGCAAAATTAAGAGAGCTGTACCAAGTATCTCAGAGCCTCCGCTCTCTATCAGT 1980
QY 2002 CCGTATGCTTCCAGAAAGCTGTGGAACGCTACTCTCTCTTTTGAGACCTTTTGGAGAGGG 2061
Db 1981 CCGTATGCTTCCAGAAAGCTGTGGAACGCTACTCTCTCTTTTGAGACCTTTTGGAGAGGG 2040
QY 2062 TCTGAAATGGTAGGCAAGAGATAGTTTCCCAAGAGATAAAAACTGGTGTGTGGCCATG 2121
Db 2041 TCTGAAATGGTAGGCAAGAGATAGTTTCCCAAGAGATAAAAACTGGTGTGTGGCCATG 2100
QY 2122 GCAGCCAAAACGGAAGGCTGAGATTCATCTCCAGAGTCCGCTCATCCACACACCCCAAT 2181
Db 2101 GCAGCCAAAACGGAAGGCTGAGATTCATCTCCAGAGTCCGCTCATCCACACACCCCAAT 2160
QY 2182 TCCAGGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTCAACATCAACGCCAGCTTCCATG 2241

Db 2161 TCAGGAGACAGCGGGAAGACATTTGCCAAGCCGGTCACCATCAGCCAGCTCCATG 2220
Qy 2242 AGGAAATCTGCACATACCTTCATAGAAAGTCCAGGAGGACTTCTGTGGTCTGAAACAC 2301
Db 2221 AGGAAATCTGCACATACCTTCATAGAAAGTCCAGGAGGACTTCTGTGGTCTGAAACAC 2280
Qy 2302 TCAACAGAATTATAGATCTTAATCTGAGTGAGTTACTGAGCTTGGTCCACATAAACACAG 2361
Db 2281 TCACAGAATTATAGATCTTAATCTGAGTGAGTTACTGAGCTTGGTCCACATAAACACAG 2340
Qy 2362 CTGAGCTTGGTCCACATAAACAGATGAAAAATACAAGAGTGACTCTATAACTCTGGTC 2421
Db 2341 CTGAGCTTGGTCCACATAAACAGATGAAAAATACAAGAGTGACTCTATAACTCTGGTC 2400
Qy 2422 TTTAAGAAAGCTGCCCTTTTCATTTTGTAGACAAAATCTTTTCAACGCTGAATGTACCTAA 2481
Db 2401 TTTAAGAAAGCTGCCCTTTTCATTTTGTAGACAAAATCTTTTCAACGCTGAATGTACCTAA 2460
Qy 2482 TCTGGTTCTACTACCATATATATATGACGTTCCCGAGGATGAATGCTGTGTTAAAT 2541
Db 2461 TCTGGTTCTACTACCATATATATATGACGTTCCCGAGGATGAATGCTGTGTTAAAT 2520
Qy 2542 TCATAAGTAAATTTGTCTACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATAT 2601
Db 2521 TCATAAGTAAATTTGTCTACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATAT 2580
Qy 2602 TCATCTTCTATATAATAGCATCCAGTTTATGGAGGCAAAAACAAGTTCTTGTTA 2661
Db 2581 TCATCTTCTATATAATAGCATCCAGTTTATGGAGGCAAAAACAAGTTCTTGTTA 2640
Qy 2662 TCCGAAACCTTTCTATGCTCAGTGGAAGTACTCCGACCCACAGCATGAGGCTGTGAA 2721
Db 2641 TCCGAAACCTTTCTATGCTCAGTGGAAGTACTCCGACCCACAGCATGAGGCTGTGAA 2700
Qy 2722 GGCTGACTGAGAAATCCTCTGCTGAAGACCCCTGGTTCTGTCTGCTCCACATGTATA 2781
Db 2701 GGCTGACTGAGAAATCCTCTGCTGAAGACCCCTGGTTCTGTCTGCTCCACATGTATA 2760
Qy 2782 ATTTTATTGAAATACATAATCTTTTCACTATG 2814
Db 2761 ATTTTATTGAAATACATAATCTTTTCACTATG 2793

RESULT 12
CS033737
LOCUS CS033737 4221 bp DNA linear PAT 10-MAR-2005
DEFINITION Sequence 3243 from Patent WO2005016962.
ACCESSION CS033737
VERSION CS033737.1 GI:60732651
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1
AUTHORS Abbas, A., Clark, H., Ouyang, W., Williams, M.P., Wood, W.I. and Wu, T.D.
TITLE Compositions and methods for the treatment of immune related diseases
JOURNAL Patent: WO 2005016962-A 3243 24-FEB-2005;
Genentech, Inc. (US)
FEATURES
source
1. .4221
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 98.5%; Score 2789.8; DB 2; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 22 CGATAACGATTTGTCTGTGAGAGCGCAACGTCGATTTCTGTGAACTTTGGAGGCATT 81
Db 1 CGATAACGATTTGTCTGTGAGAGCGCAACGTCGATTTCTGTGAACTTTGGAGGCATT 60
Qy 82 TCTACGATTTTCTCTAGCTGAGGCTTTTCTCGAGCCCTGATGCTCTTCAATTCGGTG 141
Db 61 TCTACGATTTTCTCTAGCTGAGGCTTTTCTCGAGCCCTGATGCTCTTCAATTCGGTG 120
Qy 142 CTCGCCACGCCAGCTTGGGCTCTGAGAAATGAGTGGTCTTCAATACCTCTTCAA 201
Db 121 CTCGCCACGCCAGCTTGGGCTCTGAGAAATGAGTGGTCTTCAATACCTCTTCAA 180
Qy 202 TCCCTTCTGACTGGTTATCAGTGCAAGTGAATGATGAACACACTTCTTATGGAGAAACA 261
Db 181 TCCCTTCTGACTGGTTATCAGTGCAAGTGAATGATGAACACACTTCTTATGGAGAAACA 240
Qy 262 GGAGTCCAGTTCCTCTTTTGGATGTAACCTTCTTCTGCTCCCAATATGAGAAATGTA 321
Db 241 GGAGTCCAGTTCCTCTTTTGGATGTAACCTTCTTCTGCTCCCAATATGAGAAATGTA 300
Qy 322 CTAGCAGTTGCCAATGGAAGAGGCTTTGTCGATTTGTATACACAGAAATCAAAAGTTTC 381
Db 301 CTAGCAGTTGCCAATGGAAGAGGCTTTGTCGATTTGTATACACAGAAATCAAAAGTTTC 360
Qy 382 AGAAGAAAGTCTTCAAGAAATGGATGGCTCACTGGAATGCGCTTTTGACCTGGCTGG 441
Db 361 AGAAGAAAGTCTTCAAGAAATGGATGGCTCACTGGAATGCGCTTTTGACCTGGCTGG 420
Qy 442 GTTCTGTGTAACCTTAACTTTTACAGCAGCAGGTGATCAAAACAGCCAAATTTTGGGAC 501
Db 421 GTTCTGTGTAACCTTAACTTTTACAGCAGCAGGTGATCAAAACAGCCAAATTTTGGGAC 480
Qy 502 GTTAAAGCTGTGAGCTGATTTGGACATGCAAGAGTCAATGACAGCTCAAGTCAGTT 561
Db 481 GTTAAAGCTGTGAGCTGATTTGGACATGCAAGAGTCAATGACAGCTCAAGTCAGTT 540
Qy 562 GCCTTTTCTAAGTTTGAGAAAGCTGATTTCTGTACGGTGGAAGAGATGGCAACATTATG 621
Db 541 GCCTTTTCTAAGTTTGAGAAAGCTGATTTCTGTACGGTGGAAGAGATGGCAACATTATG 600
Qy 622 GTCTGGGATACAGGTGCAACAAAAAGATGGGTTTTATAGGCAAGTGAATCAAAATCAGT 681
Db 601 GTCTGGGATACAGGTGCAACAAAAAGATGGGTTTTATAGGCAAGTGAATCAAAATCAGT 660
Qy 682 GGAGCTCAATATCCTCAGACAGCAACCCCTTCAAAACCCAGAAAGAAACAGAAATCA 741
Db 661 GGAGCTCAATATCCTCAGACAGCAACCCCTTCAAAACCCAGAAAGAAACAGAAATCA 720
Qy 742 AAAGGACTTGTCTCTCTGTGGATTTCCAGCAAAAGTGTACTGTGGTCTCTTTCAAGAC 801
Db 721 AAAGGACTTGTCTCTCTGTGGATTTCCAGCAAAAGTGTACTGTGGTCTCTTTCAAGAC 780
Qy 802 GAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAAGTATGGAAATTCAGT 861
Db 781 GAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAAGTATGGAAATTCAGT 840
Qy 862 AAGAAATATAGCTTATCGACAGAACCCATAGATCCAGTCTTCTGCTACCCAGGT 921
Db 841 AAGAAATATAGCTTATCGACAGAACCCATAGATCCAGTCTTCTGCTACCCAGGT 900
Qy 922 AGCAGCACTCGAAAACCTTGGATATTTCAAGTCTGATTTTGGATTTCCAGTGGCTCTACTTTA 981
Db 901 AGCAGCACTCGAAAACCTTGGATATTTCAAGTCTGATTTTGGATTTCCAGTGGCTCTACTTTA 960
Qy 982 TTTGCTAATTCACAGACGATAACATCTACATGTTTAAATATGATGCTGGGTTGAAGACTTCT 1041
Db 961 TTTGCTAATTCACAGACGATAACATCTACATGTTTAAATATGATGCTGGGTTGAAGACTTCT 1020
Qy 1042 CCAGTGGCTATTTCAATGGACACCAAGAACTCTACCTTTTATGTAATAATCCAGCCTTAGT 1101
Db 1021 CCAGTGGCTATTTCAATGGACACCAAGAACTCTACCTTTTATGTAATAATCCAGCCTTAGT 1080
Qy 1102 CCAGATGACCAAGTTTATGTCAGTGGCTCAAGTGAAGCTGCCTACATATGGAAGGTC 1161

```
|||||
1081 CCAGATGACCAGTCTTTAGTCACTGGCTCAAGTGATGAAGCTGCCTACATATGGAGGTC 1140
|||||
1162 TCCACACCTGGCAACCTCTACTGTCTCTGGGTCAATCTCAAGAGGTCAAGCTCTGTG 1221
|||||
1141 TCCACACCTGGCAACCTCTACTGTCTCTGGGTCAATCTCAAGAGGTCAAGCTCTGTG 1200
|||||
1222 TCCTGGTTCATCTGACTTCAAAAGATTGCTACTGTTCTGATGACAAATACACTAAAA 1281
|||||
1201 TGCTGGTTCATCTGACTTCAAAAGATTGCTACTGTTCTGATGACAAATACACTAAAA 1260
|||||
1282 ATCTGCGCTTCAATAGAGGCTTAGAGGAGAAACAGAGAGGTGATAAACTTTCCACGGTG 1341
|||||
1261 ATCTGCGCTTCAATAGAGGCTTAGAGGAGAAACAGAGAGGTGATAAACTTTCCACGGTG 1320
|||||
1342 GGTGGGCTCTCAGAGAAAAAGAGTCAAGACCTGGCTAGTAAACAGTAAACAGTAGC 1401
|||||
1321 GGTGGGCTCTCAGAGAAAAAGTCAAGACCTGGCTAGTAAACAGTAAACAGTAGC 1380
|||||
1402 CAGAGTACTCTCGCAAGCCCCCAGGGTAAAGTGCAATCCATCCAACTTTCCCGGTCA 1461
|||||
1381 CAGAGTACTCTCGCAAGCCCCCAGGGTAAAGTGCAATCCATCCAACTTTCCCGGTCA 1440
|||||
1462 TCCGAGCTTGGCCCCCAAGCTGTCTGAGACCTCCCTCTTCTTCAAACTACTCCTACG 1521
|||||
1441 TCCGAGCTTGGCCCCCAAGCTGTCTGAGACCTCCCTCTTCTTCAAACTACTCCTACG 1500
|||||
1522 TTCTCTATTAAACCTCTCTGCCAAGGCCCGGTCTCCCATCAACAGAGAGGCTCTGTC 1581
|||||
1501 TTCTCTATTAAACCTCTCTGCCAAGGCCCGGTCTCCCATCAACAGAGAGGCTCTGTC 1560
|||||
1582 TCCTCGGTCTTCCAGGCCACTTCATCTTTCAAGATGTGATTAAGAACTGGGTGACC 1641
|||||
1561 TCCTCGGTCTTCCAGGCCACTTCATCTTTCAAGATGTGATTAAGAACTGGGTGACC 1620
|||||
1642 CGAACACCTTCTCATCACCCATCATCTCCACCTGCTCGGAGACCAAGATCATGTCT 1701
|||||
1621 CGAACACCTTCTCATCACCCATCATCTCCACCTGCTCGGAGACCAAGATCATGTCT 1680
|||||
1702 CCGAGAAAGCCCTTATTCTCTGTGAGCCAGAAAGTCAATCCCAAGCAGAGGCTTGTCTGAG 1761
|||||
1681 CCGAGAAAGCCCTTATTCTCTGTGAGCCAGAAAGTCAATCCCAAGCAGAGGCTTGTCTGAG 1740
|||||
1762 TCTAGAAATAGAGTAAAGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAACAAAG 1821
|||||
1741 TCTAGAAATAGAGTAAAGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAACAAAG 1800
|||||
1822 TGCTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGATGCCAAGTTGAAATCTTCATTTG 1881
|||||
1801 TGCTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGATGGCCAAGTTGAAATCTTCATTTG 1860
|||||
1882 GATCTGTGCTTGTGCTGTGTAACAGGAAGACCTTAGTAAGGACTCTCTAGGTCCTTACC 1941
|||||
1861 GATCTGTGCTTGTGCTGTGTAACAGGAAGACCTTAGTAAGGACTCTCTAGGTCCTTACC 1920
|||||
1942 AAATCAAGCAAAATTTGAAGAGCTGGTACCAGTATCTCAGAGCCTCGCTCTCTATCATGT 2001
|||||
1921 AAATCAAGCAAAATTTGAAGAGCTGGTACCAGTATCTCAGAGCCTCGCTCTCTATCATGT 1980
|||||
2002 CCGTATGCTTCAGAAAGCTGTGAACGCTTACCTCTTCTTTCAGACCTTGTGGAGAGGG 2061
|||||
1981 CCGTATGCTTCAGAAAGCTGTGAACGCTTACCTCTTCTTTCAGACCTTGTGGAGAGGG 2040
|||||
2062 TCTGAAATGGTAGGCAAGAGATAGTTCCCGAGAGAAATAAAACTGGTTGTGGCCATG 2121
|||||
2041 TCTGAAATGGTAGGCAAGAGATAGTTCCCGAGAGAAATAAAACTGGTTGTGGCCATG 2100
|||||
2122 GCAGCAAAACGGAAGCTGAGAAATCCATCTCAGAAAGTCCGTCTATCCGAGACCCCAAT 2181
|||||
2101 GCAGCAAAACGGAAGCTGAGAAATCCATCTCAGAAAGTCCGTCTATCCGAGACCCCAAT 2160
|||||
2182 TCCAGGAGACAGCGGAAAGACATTGGCAAGCCCGGTCCACCATCAGCCCGAGCTTCCATG 2241
|||||
```

```
Db 2161 TCCAGGAGACAGAGCGGAAAGACATTCGCAAGCCCGGTGCACCATCAACGCCAGCTCCATG 2220
Qy 2242 AGGAAATCTGCAATATCTTCCATAGAAAGTCCGAGGAGACTTCTGTGGTCTTGAACAC 2301
Db 2221 AGGAAATCTGCAATATCTTCCATAGAAAGTCCGAGGAGACTTCTGTGGTCTTGAACAC 2280
Qy 2302 TCAACAGAAATATAGATTCTTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTTAAACAAAG 2361
Db 2281 TCAACAGAAATATAGATTCTTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTTAAACAAAG 2340
Qy 2362 CTGAGCTTTGGTCCACTTAAACAGATGAAAAATACAGAGTACTCTATAACTCTGGTC 2421
Db 2341 CTGAGCTTTGGTCCACTTAAACAGATGAAAAATACAGAGTACTCTATAACTCTGGTC 2400
Qy 2422 TTTAAGAAAGCTGCTTTTCAATTTTATAGACAAATCTTTTCAACGCTGAAATGTACCTAA 2481
Db 2401 TTTAAGAAAGCTGCTTTTCAATTTTATAGACAAATCTTTTCAACGCTGAAATGTACCTAA 2460
Qy 2482 TCTGGTCTTACTACCAATATGATATGACGCTTCCGAGGATGAATGCTGTGTTTAAATT 2541
Db 2461 TCTGGTCTTACTACCAATATGATATGACGCTTCCGAGGATGAATGCTGTGTTTAAATT 2520
Qy 2542 TCATAAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTTAT 2601
Db 2521 TCATAAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTTAT 2580
Qy 2602 TCATCTTCTCTATAATTAATGACATCCAGTTTCATGGAGGCAAAACAGAGTTCTTGTTA 2661
Db 2581 TCATCTTCTCTATAATTAATGACATCCAGTTTCATGGAGGCAAAACAGAGTTCTTGTTA 2640
Qy 2662 TCCGAAACTTTCTATGCTCAGTGGAAGTATCTGCCAGCCACAGCATGAGGCTGTGAA 2721
Db 2641 TCCGAAACTTTCTATGCTCAGTGGAAGTATCTGCCAGCCACAGCATGAGGCTGTGAA 2700
Qy 2722 GGCTGACTGGAATCTCTGCTGAAGACCCCTGGTTCTGTTCTGCTCCCAACATGTATA 2781
Db 2701 GGCTGACTGGAATCTCTGCTGAAGACCCCTGGTTCTGTTCTGCTCCCAACATGTATA 2760
Qy 2782 ATTTTATTTGAAATACATAATCTTTTCACTATG 2814
Db 2761 ATTTTATTTGAAATACATAATCTTTTCACTATG 2793

RESULT 13
CS042689 LOCUS CS042689 4221 bp DNA linear PAT 22-MAR-2005
DEFINITION Sequence 3243 from Patent WO2005019258.
ACCESSION CS042689
VERSION CS042689.1 GI:61849834
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1
REFERENCE
AUTHORS Abbas,A., Clark,H., Ouyang,W., Williams,P.M., Wood,W.I. and Wu,T.D.
TITLE Compositions and methods for the treatment of immune related
diseases
JOURNAL Patent: WO 2005019258-A 3243 03-MAR-2005;
Genentech, Inc. (US)
FEATURES
source Location/Qualifiers
1..4221
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 98.5%; Score 2789.8; DB 2; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 22 CGATAACGATTGTGTTGTGAGAGCGCAACGTCGATTTTCTGCTGAACCTTGGAGGCATT 81
```

Db	1	CGATAACGATTTGTGTGTGAGAGGCGCAAGCTCGGATTTCTGCTGAACTTGGAGGCAATT	60
Qy	82	TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAAATTCGGTG	141
Db	61	TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAAATTCGGTG	120
Qy	142	CTCCGCCAGGCCCAAGCTTGGCGTCTCAGAAATGGATGGTCTTCAAAATACCCCTCTTCAA	201
Db	121	CTCCGCCAGGCCCAAGCTTGGCGTCTCAGAAATGGATGGTCTTCAAAATACCCCTCTTCAA	180
Qy	202	TCCCTTTCTGACTGGTTATCAGTGCAAGTGGTAATGATGAACACACTTCTTATGGAGAAACA	261
Db	181	TCCCTTTCTGACTGGTTATCAGTGCAAGTGGTAATGATGAACACACTTCTTATGGAGAAACA	240
Qy	262	GGAGTCCCAAGTTCTCTCTTTTGGATGTACCTTCTTCTGCTGCCAATATGGAACATGTA	321
Db	241	GGAGTCCCAAGTTCTCTCTTTTGGATGTACCTTCTTCTGCTGCCAATATGGAACATGTA	300
Qy	322	CTAGCAGTTGCAATGAAGAGGCTTTGTTGCAATGTATGAACAGAAATCAAAAGTTTC	381
Db	301	CTAGCAGTTGCAATGAAGAGGCTTTGTTGCAATGTATGAACAGAAATCAAAAGTTTC	360
Qy	382	AGAAAGATGTCTTCAAAGAAATGGAATGGCTCACCTGGAATGCCGTCTTGGACCTGGCTGG	441
Db	361	AGAAAGATGTCTTCAAAGAAATGGAATGGCTCACCTGGAATGCCGTCTTGGACCTGGCTGG	420
Qy	442	GTTCTCTGCTGAACCTTAAACCTTTGTACAGCAGCAGTGATCAAAACAGCCAAATTTTGGAC	501
Db	421	GTTCTCTGCTGAACCTTAAACCTTTGTACAGCAGCAGTGATCAAAACAGCCAAATTTTGGAC	480
Qy	502	GTAAGAGCTGGTGAAGCTGATTTGGAAACATGCAAAAGTCAATCAATGCAGCTCAAGTCAGTT	561
Db	481	GTAAGAGCTGGTGAAGCTGATTTGGAAACATGCAAAAGTCAATCAATGCAGCTCAAGTCAGTT	540
Qy	562	GCCTTTTCTAAGTTTGAAAGCTGTATTTCTGTACGGGTGGAAGAGATGGCAACATTTATG	621
Db	541	GCCTTTTCTAAGTTTGAAAGCTGTATTTCTGTACGGGTGGAAGAGATGGCAACATTTATG	600
Qy	622	GTCTGGGATACCAAGTCCAAACAAAAGATGGGTTTATAGGCAAGTGAATCAAAATCAGT	681
Db	601	GTCTGGGATACCAAGTCCAAACAAAAGATGGGTTTATAGGCAAGTGAATCAAAATCAGT	660
Qy	682	GGAGCTCACAATACCTCAGACAAGAAACCCCTTCAAAAACCAAGAAAGAAACAGAAATCA	741
Db	661	GGAGCTCACAATACCTCAGACAAGAAACCCCTTCAAAAACCAAGAAAGAAACAGAAATCA	720
Qy	742	AAAGGACTTGCTCTCTGTGGATTTCCAGAAAGTGTACTGTGGTCTCTTTCAAGAC	801
Db	721	AAAGGACTTGCTCTCTGTGGATTTCCAGAAAGTGTACTGTGGTCTCTTTCAAGAC	780
Qy	802	GAGNAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATATCAAGATATGGGATTTACGT	861
Db	781	GAGNAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATATCAAGATATGGGATTTACGT	840
Qy	862	AAGAAATATATCTGCTTATCGACAAGAACCCATAGCATCCCAAGTCTTTCCCTGTACCCAGGT	921
Db	841	AAGAAATATATCTGCTTATCGACAAGAACCCATAGCATCCCAAGTCTTTCCCTGTACCCAGGT	900
Qy	922	AGCAGCACTCGAAAACCTTGGATATTCAGTCTGATTTTGGATTTCCACTGGCTCTACTTTTA	981
Db	901	AGCAGCACTCGAAAACCTTGGATATTCAGTCTGATTTTGGATTTCCACTGGCTCTACTTTTA	960
Qy	982	TTTGCTAATTTGCACAGACGATAACATCTACATGTTTTAATATGACTGGGTGAAGACTTCT	1041
Db	961	TTTGCTAATTTGCACAGACGATAACATCTACATGTTTTAATATGACTGGGTGAAGACTTCT	1020
Qy	1042	CCAGTGGCTATTTTCAATGGACACCAAGACTCTACCTTTTATGTGTAATCCAGCTTAGT	1101
Db	1021	CCAGTGGCTATTTTCAATGGACACCAAGACTCTACCTTTTATGTGTAATCCAGCTTAGT	1080
Qy	1102	CCAGATGACCAAGTTTTTTAGTCAGTGGCTCAAGTGATGAAGCTGCTACATATGGAAGTTC	1161

1081	DB	CCAGATGACCAAGTTTTTAGTTCAGTGGCTCAAGTGTGTAAGCTGCTACATATATGAAGGTC	1144
1162	QY	TCCACACCCTGGCAACCTCCTACTGTGCTCCTGGGTCAITTTCTCAAGAGGTCACGTCTGTG	1221
1141	DB	TCCACACCCTGGCAACCTCCTACTGTGCTCCTGGGTCAITTTCTCAAGAGGTCACGTCTGTG	1200
1222	QY	TGCTGGTGTCCATCTGACTTCACAAAGATTGCTACCTGTTCTGTGATGCAATACACTTAAAA	1281
1201	DB	TGCTGGTGTCCATCTGACTTCACAAAGATTGCTACCTGTTCTGTGATGCAATACACTTAAAA	1260
1282	QY	ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACCAAGAGGTGATAAATTTTCCACGGTG	1341
1261	DB	ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACCAAGAGGTGATAAATTTTCCACGGTG	1320
1342	QY	GGTTGGGCGCTTCAGAAAGAAAAAGAGTCAAGACCTGGCCTAGTAAACAGTAACGAGTAGC	1401
1321	DB	GGTTGGGCGCTTCAGAAAGAAAAAGAGTCAAGACCTGGCCTAGTAAACAGTAACGAGTAGC	1380
1402	QY	CAGAGTACTCTGTCCAAAGCCCCCAGGGTAAAGTGCAATCCATCCAATTTCTTCCCGTCA	1461
1381	DB	CAGAGTACTCTGTCCAAAGCCCCCAGGGTAAAGTGCAATCCATCCAATTTCTTCCCGTCA	1440
1462	QY	TCCGACGCTGTGCCCCCAAGCTGTCTGCGAGACCTCCCTCTTCTTCAAAATACCTCTTACG	1521
1441	DB	TCCGACGCTGTGCCCCCAAGCTGTCTGCGAGACCTCCCTCTTCTTCAAAATACCTCTTACG	1500
1522	QY	TTCTCTATTAAACCTCTCTGTCCAAAGCCCCGGTCTCCCATCAACAGAGAGGCTCTGTGTC	1581
1501	DB	TTCTCTATTAAACCTCTCTGTCCAAAGCCCCGGTCTCCCATCAACAGAGAGGCTCTGTGTC	1560
1582	QY	TCCTCCGCTCTCTCCAAAGCCACTTTCAATCTTTCAAAGATGTCGATTAGAAACTGTGGTGACC	1641
1561	DB	TCCTCCGCTCTCTCCAAAGCCACTTTCAATCTTTCAAAGATGTCGATTAGAAACTGTGGTGACC	1620
1642	QY	CGAACACCTTCTCATCACCACCATCACCTCCACCTGTTCCGAGACCAAGATCATCTGCT	1701
1621	DB	CGAACACCTTCTCATCACCACCATCACCTCCACCTGTTCCGAGACCAAGATCATCTGCT	1680
1702	QY	CCGAGAAAAGCCCTATTTCCTGTGAGCCAGAAGTCATCCCAAGCAGAGGCTTGCTCTGAG	1761
1681	DB	CCGAGAAAAGCCCTATTTCCTGTGAGCCAGAAGTCATCCCAAGCAGAGGCTTGCTCTGAG	1740
1762	QY	TCTAGAAATAGAGTAAAGAGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAAG	1821
1741	DB	TCTAGAAATAGAGTAAAGAGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAAG	1800
1822	QY	TGTTGAAGAGTTGTTAACTGTGTGACTGAGCTTGATGGCCAAAGTTGAAATCTTCATTG	1881
1801	DB	TGTTGAAGAGTTGTTAACTGTGTGACTGAGCTTGATGGCCAAAGTTGAAATCTTCATTG	1860
1882	QY	GATCTGTGCTGCTTGTCTGGTAAACAGAGAGACCTTAGTAGGACTCTCTAGGTCCTTACC	1941
1861	DB	GATCTGTGCTGCTTGTCTGGTAAACAGAGAGACCTTAGTAGGACTCTCTAGGTCCTTACC	1920
1942	QY	AAATCAAGCAAAATTGAAGGAGCTGGTACCAGTATCTCAGAGCCTCCGCTCTCTATCAGT	2001
1921	DB	AAATCAAGCAAAATTGAAGGAGCTGGTACCAGTATCTCAGAGCCTCCGCTCTCTATCAGT	1980
2002	QY	CCGTATGCTTCAGAAAAGCTGTGGAAACGCTACTCTTCTTTGAGACCTTGTGGAGAAAGG	2061
1981	DB	CCGTATGCTTCAGAAAAGCTGTGGAAACGCTACTCTTCTTTGAGACCTTGTGGAGAAAGG	2040
2062	QY	TCTGAAATGGTGGCAAGAGATAGTTTCCCGAGAGATAAAACCTGGTTGTTGGCCATG	2121
2041	DB	TCTGAAATGGTGGCAAGAGATAGTTTCCCGAGAGATAAAACCTGGTTGTTGGCCATG	2100
2122	QY	GCAGCCAAACGGAAAGGCTGGAATTCATCTCCAGAAAGTCCGCTCATCCAGACACCCAAT	2181
2101	DB	GCAGCCAAACGGAAAGGCTGGAATTCATCTCCAGAAAGTCCGCTCATCCAGACACCCAAT	2160
2182	QY	TCCAGAGACAGAGCGGAAAGACATTTGCCAAGCCCCGCTCACATCAGCCCCAGCTCCATG	2241
2161	DB	TCCAGAGACAGAGCGGAAAGACATTTGCCAAGCCCCGCTCACATCAGCCCCAGCTCCATG	2220

QY	2242	AGGAAATCTGCACATACCTTCCATAGAAAGTCCCAAGGAGACTTCTGTGGTCTCGAACAC	2301
Db	2221	AGGAAATCTGCACATACCTTCCATAGAAAGTCCCAAGGAGACTTCTGTGGTCTCGAACAC	2280
QY	2302	TCAACAGAAATATAGATCTTAATCTGAGTGAAGTAACTGAGCTTTGGTCCACCTAAACAAAG	2361
Db	2281	TCAACAGAAATATAGATCTTAATCTGAGTGAAGTAACTGAGCTTTGGTCCACCTAAACAAAG	2340
QY	2362	CTGAGCTTTGGTCCACTAAACCAAGATGAAAAATACAAGAGTGACTCTATAACTCTGGTC	2421
Db	2341	CTGAGCTTTGGTCCACTAAACCAAGATGAAAAATACAAGAGTGACTCTATAACTCTGGTC	2400
QY	2422	TTTAAGAAAGTGCCTTTTCATTTTATAGACAAATCTTTTCAAGCTGAAATGACCTAA	2481
Db	2401	TTTAAGAAAGTGCCTTTTCATTTTATAGACAAATCTTTTCAAGCTGAAATGACCTAA	2460
QY	2482	TCGTGTTCTACTACCATATATGATGATGAGCTTCCGAGGAGATGAATGCTGTGTTTAAATT	2541
Db	2461	TCGTGTTCTACTACCATATATGATGATGAGCTTCCGAGGAGATGAATGCTGTGTTTAAATT	2520
QY	2542	TCATAAGTAAATTTGTCACTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTTAT	2601
Db	2521	TCATAAGTAAATTTGTCACTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTTAT	2580
QY	2602	TCATCTTCTATATAATAGATCCAGTTCATGAGGAGCAAAAAACAAGTTCTTGTTA	2661
Db	2581	TCATCTTCTATATAATAGATCCAGTTCATGAGGAGCAAAAAACAAGTTCTTGTTA	2640
QY	2662	TCCTGAACTTCTATGCTCAGTGAAGATCTCCAGGACACACATGAGGCTGTGAA	2721
Db	2641	TCCTGAACTTCTATGCTCAGTGAAGATCTCCAGGACACACATGAGGCTGTGAA	2700
QY	2722	GGCTGACTGAGAAATCTCTGCTGAAGACCCCTGTTCTGTTCTGCTCCCAACATGATA	2781
Db	2701	GGCTGACTGAGAAATCTCTGCTGAAGACCCCTGTTCTGTTCTGCTCCCAACATGATA	2760
QY	2782	ATTTATTTGAAATACATAATCTTTTCACTATG	2814
Db	2761	ATTTATTTGAAATACATAATCTTTTCACTATG	2793
RESULT 14			
AF195765			
LOCUS			
DEFINITION	Homo sapiens L2DTL protein (L2DTL) mRNA, linear PRI 21-FEB-2000		
ACCESSION	AF195765		
VERSION	AF195765.1 GI:7012713		
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Mueller, R. and Ziegler, B.L.		
TITLE	Identification of L2DTL, a human WD-40 repeat gene homolog of the Drosophila lethal (2) denticleless heat shock gene [1(2)dtl]		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 4221)		
AUTHORS	Mueller, R. and Ziegler, B.L.		
TITLE	Direct Submission		
JOURNAL	Submitted (18-OCT-1999) Kimmel Cancer Institute, Thomas Jefferson University, 233 South 10th Street, Philadelphia, PA 19107-5541, USA		
FEATURES	Location/Qualifiers		
source	1..4221		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
gene	1..4221		
	/gene="L2DTL"		
CDS	103..2295		
	/gene="L2DTL"		

/note="contains WD-40 repeats; similar to the Drosophila lethal (2) denticleless heat shock gene, 1(2)dtl"
/codon_start=1
/product="L2DTL protein"
/protein_id="AAF35182.1"
/db_xref="GI:7012714"
/translation="MLPNSVLROPLQGLVRLNGWSSOYPLQSLLTGYOCSGNDEHTSVG
ETGVVPFGCTFTSSAPNMEHVLAVANEEGFVRLYNTESSQSPKPKKEWHAHNAVF
DLAWPGLKLVTAAGDQTAFAKFDWVAGELIGTCKGHQCSLSKSVAFKFAVFCVGG
RDGNIWMDTRCNKKGDFYQVQNIQSGAHTSDKQTPSPKPKKQNSGLAPVDFQGS
VTVLFDENTLVSAGAVDGIKVMDFLRKNYTAIRQEPFIASKSLYFGSSTRKLGYGS
LILDSGTSTIPANCTDDNIYMFNMTGLKTSFPAIFNGHONSTFVFKSLSPDDQFLVS
GSDRAAYIKVSTPMOPPTVLLGHSEVTSVCMCPDSDFTKIATCSDNLTIKIWLNR
GLBEKPGDKLSLVGHWASOKKESRGLVTVTSOSTPAKAPRVKCPNPSNPSAAAC
APSCAGLPPLPSNTPTFSIKTISPAKARSPINRRGSVSVSPKPSFPMIRNWWTRT
PSSSPKTPPASSETKIMSPKALIPVSKSQBAECSESRNRVRRVRLDSSCLSEVKOK
CVKSCNCTBELDQVENLHLDCLCAGNQLDLSKSLGPTKSKIEGAGTSISPPSP
ISPVASRSCGTLPPLPRPCGEGEMVGKNSPENKWLMAAKRAENPSPRSPSS
QTPNSRRQSGKTLPSPV:ITPSSMRKICTYFHRKSQEDFCGPEHSTEL"

ORIGIN

Query Match		98.5%;	Score 2789.8;	DB 5;	Length 4221;
Best Local Similarity		99.9%;	Pred. No. 0;		
Matches 2791;		Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY	22	CGATAACGATTTGTGTTGTGAGAGCGCCAGCTGCGATTTCTGCTGAACCTTGGAGGCATT	81		
Db	1	CGATAACGATTTGTGTTGTGAGAGCGCCAGCTGCGATTTCTGCTGAACCTTGGAGGCATT	60		
QY	82	TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG	141		
Db	61	TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG	120		
QY	142	CTCCGACGACCCCGAGCTTGGCGTCTTGAGAAATGGATGGTCTTCAATAACCCCTTCAA	201		
Db	121	CTCCGACGACCCCGAGCTTGGCGTCTTGAGAAATGGATGGTCTTCAATAACCCCTTCAA	180		
QY	202	TCCCTTCTGACTGGTTATCAGTCAGTCAGTGGTAAATGATGAACACACTTCTTATGGAGAAC	261		
Db	181	TCCCTTCTGACTGGTTATCAGTCAGTCAGTGGTAAATGATGAACACACTTCTTATGGAGAAC	240		
QY	262	GGAGTCCCAAGTTCTCTCTTTTGGATGTACCTTCTCTGCTCCCAATATGGAACATGTA	321		
Db	241	GGAGTCCCAAGTTCTCTCTTTTGGATGTACCTTCTCTGCTCCCAATATGGAACATGTA	300		
QY	322	CTAGCAGTTGCCAATGAAGAAGCTTTGTTTCGATTTGTATATAACACAGAAATCAAAAGTTTC	381		
Db	301	CTAGCAGTTGCCAATGAAGAAGCTTTGTTTCGATTTGTATATAACACAGAAATCAAAAGTTTC	360		
QY	382	AGAAAGAGTGTCTTCAAAGATGGATGGCTCAGTCAGTGGTAAATGATGAACACACTTCTTATGGAGAAC	441		
Db	361	AGAAAGAGTGTCTTCAAAGATGGATGGCTCAGTCAGTGGTAAATGATGAACACACTTCTTATGGAGAAC	420		
QY	442	GTTCTGTGTAACCTTAAACCTTTGTACAGCAGCAGGTGATCAAAACAGCCAAATTTTGGGAC	501		
Db	421	GTTCTGTGTAACCTTAAACCTTTGTACAGCAGCAGGTGATCAAAACAGCCAAATTTTGGGAC	480		
QY	502	GTAAGAGCTGGTGAAGCTGATTTGGAACATGCAAGAGTCAATCAATCAGCCTCAAGTCAGTT	561		
Db	481	GTAAGAGCTGGTGAAGCTGATTTGGAACATGCAAGAGTCAATCAATCAGCCTCAAGTCAGTT	540		
QY	562	GCCTTTTCTAAGTTTTCAGAAAGCTGTATTCTGTACGGGTGGAAGAGATGGCAACATTATG	621		
Db	541	GCCTTTTCTAAGTTTTCAGAAAGCTGTATTCTGTACGGGTGGAAGAGATGGCAACATTATG	600		
QY	622	GTCTGGGATACCAAGTGCACAAACAAAAGATGGGTTTTATAGGCAAGTGAATCAAAATCAGT	681		
Db	601	GTCTGGGATACCAAGTGCACAAACAAAAGATGGGTTTTATAGGCAAGTGAATCAAAATCAGT	660		
QY	682	GGAGCTCAATAATCCTCAGCAAGCAAAACCCCTTCAAAACCCCAAGAGAAACAGAAATTC	741		
Db	661	GGAGCTCAATAATCCTCAGCAAGCAAAACCCCTTCAAAACCCCAAGAGAAACAGAAATTC	720		

QY 742 AAAGACTTGTCTCTCTGTGGATTTCCAGCAAAAGTTGTTACTGTGTCTCTTTCAAGAC 801
DB |||||
QY 721 AAAGGACTTGTCTCTCTGTGGATTTCCAGCAAAAGTTGTTACTGTGTCTCTTTCAAGAC 780
DB |||||
QY 802 GAGATACCTTAGTCTCAGCAGAGAGCTGTGGATGGGATAATCAAGATATGGGATTTACGT 861
DB |||||
QY 781 GAGATACCTTAGTCTCAGCAGAGAGCTGTGGATGGGATAATCAAGATATGGGATTTACGT 840
DB |||||
QY 862 AAGAAATTACTGTCTTATCGACAGAAACCATAGCATCCAAAGTCTTTCTGTGTACCCAGGT 921
DB |||||
QY 841 AAGAAATTACTGTCTTATCGACAGAAACCATAGCATCCAAAGTCTTTCTGTGTACCCAGGT 900
DB |||||
QY 922 AGCAGCACTCGAAAACCTTGGATATTCAGTCTGATTTTGGATTTCCACTGGCTCTACTTTTA 981
DB |||||
QY 901 AGCAGCACTCGAAAACCTTGGATATTCAGTCTGATTTTGGATTTCCACTGGCTCTACTTTTA 960
DB |||||
QY 982 TTTGCTAATTGACAGACGATTAACATCTACATGTTTAAATATGATGCTGGGTGGAAGACTTCT 1041
DB |||||
QY 961 TTTGCTAATTGACAGACGATTAACATCTACATGTTTAAATATGATGCTGGGTGGAAGACTTCT 1020
DB |||||
QY 1042 CCAGTGGCTATTTTCAATGCGACACCAAGAACTCTACCTTTTATGTAAATCCAGCCCTTAGT 1101
DB |||||
QY 1021 CCAGTGGCTATTTTCAATGCGACACCAAGAACTCTACCTTTTATGTAAATCCAGCCCTTAGT 1080
DB |||||
QY 1102 CCAGATGACCAAGTTTGTAGTCAAGTGGCTCAAGTGAATGAAGCTGCCTACATATGGAAGGTC 1161
DB |||||
QY 1081 CCAGATGACCAAGTTTGTAGTCAAGTGGCTCAAGTGAATGAAGCTGCCTACATATGGAAGGTC 1140
DB |||||
QY 1162 TCACACCTCTGGCAACCTCTACTGTGCTCTCTGGGTCAATTCCTCAAGAGATCAAGTCTGTG 1221
DB |||||
QY 1141 TCACACCTCTGGCAACCTCTACTGTGCTCTCTGGGTCAATTCCTCAAGAGGTCACGCTGTG 1200
DB |||||
QY 1222 TGCTGGTGTCCATCTGACTTCCAAAGATTGCTACTGTGTTCTGTGATGACAAATACACTAAA 1281
DB |||||
QY 1201 TGCTGGTGTCCATCTGACTTCCAAAGATTGCTACTGTGTTCTGTGATGACAAATACACTAAA 1260
DB |||||
QY 1282 ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGAGAGGTGATAAAACCTTTCCACGGTG 1341
DB |||||
QY 1261 ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGAGAGGTGATAAAACCTTTCCACGGTG 1320
DB |||||
QY 1342 GGTGGGCTCTCAGAAAGAAAAGAGTCAAGACCTGGGCTAGTAAACGATAAACGAGTAGC 1401
DB |||||
QY 1321 GGTGGGCTCTCAGAAAGAAAAGAGTCAAGACCTGGGCTAGTAAACGATAAACGAGTAGC 1380
DB |||||
QY 1402 CAGAGTACTCTCCCAAGGCCCCAGGTTAAAGTGCATCCATCCAAATCTTTCCCGGTCA 1461
DB |||||
QY 1381 CAGAGTACTCTCCCAAGGCCCCAGGTTAAAGTGCATCCATCCAAATCTTTCCCGGTCA 1440
DB |||||
QY 1462 TCCGAGCTTGTGCCCCAAGCTGTGTGGAGACCTCCCTCTTCTTCAATATCTCTACG 1521
DB |||||
QY 1441 TCCGAGCTTGTGCCCCAAGCTGTGTGGAGACCTCCCTCTTCTTCAATATCTCTACG 1500
DB |||||
QY 1522 TTCTCTATTAAAACCTCTCTCCCAAGGCCCCGGTCTCCCATCAACAGAAAGAGGCTCTGTG 1581
DB |||||
QY 1501 TTCTCTATTAAAACCTCTCTCCCAAGGCCCCGGTCTCCCATCAACAGAAAGAGGCTCTGTG 1560
DB |||||
QY 1582 TCCTCGTCTCTCCCAAGGCCCCGCTTCTTCAAGATGTCGATTAGAAACTGGGTGACC 1641
DB |||||
QY 1561 TCCTCGTCTCTCCCAAGGCCCCGCTTCTTCAAGATGTCGATTAGAAACTGGGTGACC 1620
DB |||||
QY 1642 CGAAACACCTTCTCATCACACCATCTCTCAACCTGTTCGAGACCAAGATCATGTCT 1701
DB |||||
QY 1621 CGAAACACCTTCTCATCACACCATCTCTCAACCTGTTCGAGACCAAGATCATGTCT 1680
DB |||||
QY 1702 CCGAGAAAAGCCCTTATTCTGTGAGCCAGAGTCAATCCCAAGCAGAGGCTTGTCTGTGAG 1761
DB |||||
QY 1681 CCGAGAAAAGCCCTTATTCTGTGAGCCAGAGTCAATCCCAAGCAGAGGCTTGTCTGTGAG 1740
DB |||||
QY 1762 TCTAGAAATAGAGTAAAGAGGAGGTAGACTCAAGCTGTCTGGAGAGTGTGAAAACAAAG 1821
DB |||||
QY 1741 TCTAGAAATAGAGTAAAGAGGAGGTAGACTCAAGCTGTCTGGAGAGTGTGAAAACAAAG 1800
DB |||||
QY 1822 TGTGTGAAGAGTTGTAACCTGTGACTGAGCTTGATGGCCAAAGTTGAAATCTTCAATTG 1881
DB |||||

DB 1801 TGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGATGCCAAGTTGAAAACTTCTCATTTG 1860
QY |||||
DB 1882 GATCTGTCTGCTGCTGTGTTAAACAGGAAGACCTTAGTAGGACTCTCTAGGTCCTTACC 1941
DB |||||
DB 1861 GATCTGTCTGCTGCTGTGTTAAACAGGAAGACCTTAGTAGGACTCTCTAGGTCCTTACC 1920
QY |||||
DB 1942 AAATCAAGCAAAATTTGAAGGAGCTGGTACAGTAGTATCTCAGAGCCCTCCGCTCTCTATCAGT 2001
DB |||||
DB 1921 AAATCAAGCAAAATTTGAAGGAGCTGGTACAGTAGTATCTCAGAGCCCTCCGCTCTCTATCAGT 1980
QY |||||
QY 2002 CCGTATGCTTTCAGAAAGCTGTGGAACGCTACTCTTCTTTTGAAGACCTTTGTGGAGAAAGG 2061
DB |||||
DB 1981 CCGTATGCTTTCAGAAAGCTGTGGAACGCTACTCTTCTTCTTTGAGACCTTTGTGGAGAAAGG 2040
QY |||||
QY 2062 TCTGAAATGCTTAGGCAAGAGAAATAGTTCCCGAGAGAAATAAACTGGTGTGTGGCCATG 2121
DB |||||
DB 2041 TCTGAAATGCTTAGGCAAGAGAAATAGTTTCCCGAGAGAAATAAACTGGTGTGTGGCCATG 2100
QY |||||
QY 2122 GCAGCAAAACGGAAGGCTGAGAAATCCATCTCCAGAAAGTCCGCTCATCCAGACACCCCAAT 2181
DB |||||
DB 2101 GCAGCAAAACGGAAGGCTGAGAAATCCATCTCCAGAAAGTCCGCTCATCCAGACACCCCAAT 2160
QY |||||
DB 2182 TCCAGGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTCAACATCAACGCCAGCTTCCATG 2241
DB |||||
DB 2161 TCCAGGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTCAACATCAACGCCAGCTTCCATG 2220
QY |||||
QY 2242 AGGAAATCTGCACATACTTCCATAGAAAGTCCCGAGGAGGACTTCTGTGGTCTCTGAAAC 2301
DB |||||
DB 2221 AGGAAATCTGCACATACTTCCATAGAAAGTCCCGAGGAGGACTTCTGTGGTCTCTGAAAC 2280
QY |||||
QY 2302 TCAACAGAAATTTATAGATTCTTAATCTGAGTGTGTTACTCAGCTTTTGGTCCACTTAAACAAAG 2361
DB |||||
DB 2281 TCAACAGAAATTTATAGATTCTTAATCTGAGTGTGTTACTGAGCTTTGGTCCACTTAAACAAAG 2340
QY |||||
QY 2362 CTGAGCTTTTGGTCCACTTAAACCAAGATGAAAAATAACAAGAGTGACTCTTATAACTCTGTGTC 2421
DB |||||
DB 2341 CTGAGCTTTTGGTCCACTTAAACCAAGATGAAAAATAACAAGAGTGACTCTTATAACTCTGTGTC 2400
QY |||||
QY 2422 TTTAAGAAAGCTGCTTTTCAATTTTATAGACAAATCTTTTCAACGCTGAAATGTACTTAA 2481
DB |||||
DB 2401 TTTAAGAAAGCTGCTTTTCAATTTTATAGACAAATCTTTTCAACGCTGAAATGTACTTAA 2460
QY |||||
QY 2482 TCTGTCTCTACTACCATTAATGTATATGAGCTTCCGAGGAGTGAATGTGTGTTTAAAT 2541
DB |||||
DB 2461 TCTGTCTCTACTACCATTAATGTATATGAGCTTCCGAGGAGTGAATGTGTGTTTAAAT 2520
QY |||||
QY 2542 TCATAAAGTAAATTTGTCTACTCTAGCAATTTTGAATGAATGATCTTCACTTTTAAATTTAT 2601
DB |||||
DB 2521 TCATAAAGTAAATTTGTCTACTCTAGCAATTTTGAATGAATGATCTTCACTTTTAAATTTAT 2580
QY |||||
QY 2602 TCATCTCTCTATAATAATATGATCCCGAGTTCATGGAGGCAAAAACAAGTTTCTTGTTA 2661
DB |||||
DB 2581 TCATCTCTCTATAATAATATGATCCCGAGTTCATGGAGGCAAAAACAAGTTTCTTGTTA 2640
QY |||||
QY 2662 TCCTGAAATCTTCTATGCTCAGTGTGAAAGTATCTGCCAGCCACAGCATGAGGCTGTGAA 2721
DB |||||
DB 2641 TCCTGAAATCTTCTATGCTCAGTGTGAAAGTATCTGCCAGCCACAGCATGAGGCTGTGAA 2700
QY |||||
QY 2722 GGCTGACTGAGAAATCTCTCTGCTGAAAGACCCCTGGTCTGTGCTTCCCTTCAACATGTATA 2781
DB |||||
DB 2701 GGCTGACTGAGAAATCTCTCTGCTGAAAGACCCCTGGTCTGTGCTTCCCTTCAACATGTATA 2760
QY |||||
QY 2782 ATTTTATTTGAAATACATAAATCTTTTCACTATG 2814
DB |||||
DB 2761 ATTTTATTTGAAATACATAAATCTTTTCACTATG 2793
DB |||||

RESULT 15
BC033297
LOCUS
DEFINITION
BC033297
Homo sapiens denticleless homolog (Drosophila), mRNA (cDNA clone
MGC:24863 IMAGE:4342989), complete cds.

```
ACCESSION BC033297
VERSION BC033297.1 GI:23271027
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 2782)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheet,T.E., Brownstein,M.J., Ustin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Greenwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalios,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2782)
AUTHORS NIH MGC Project
Direct Submission
Submitted (26-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue procurement: Lou Staudt
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louleeged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Place: 29 Row: a Column: 24
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 7705575.
FEATURES
Location/Qualifiers
1..2782
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:24863 IMAGE:4342989"
/tissue_type="Lymph, lymphoma"
/clone_lib="NIH MGC_85"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
1..2782
/gene="DTL"
/note="synonyms: L2DTL, RAMP"
/db_xref="GeneID:51514"
80..2272
gene
CDS
```


This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2006, 05:50:06 ; Search time 1688 Seconds

(without alignments)
11693.386 Million cell updates/sec

Title: US-10-726-160-1

Perfect score: 2831

Sequence: 1 ggcacgagcgaggttgag.....atgaaaaaaaaaaaaaaaaa 2831

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5244920 seqs, 3486124231 residues

Total number of hits satisfying chosen parameters: 10489840

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 8:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*
- 14: geneseqn2005s:*
- 15: geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	2831	100.0	2831 10	Adf69739 Gene enco
2	2831	100.0	2831 14	Adx69660 Human ret
3	2816.2	99.5	2838 4	Aah99153 Human pro
4	2804	99.0	4422 4	Aah72681 Human cer
5	2804	99.0	4422 5	Abv25710 Human pro
6	2804	99.0	4422 5	Adl61974 Human ova
7	2802	99.0	4450 10	Adf82011 Leukaemia
8	2789.8	98.5	4221 10	Aad63212 Human DNA
9	2789.8	98.5	4221 10	Aad62813 Human DNA
10	2789.8	98.5	4221 10	Abx95498 Human ret
11	2789.8	98.5	4221 11	Adp88308 Metastati
12	2789.8	98.5	4221 12	Adp07299 Human L2D
13	2789.8	98.5	4221 12	Ado20068 Human PRO
14	2789.8	98.5	4221 12	Ado20231 Human PRO
15	2789.8	98.5	4221 13	Adr25784 Breast ca
16	2789.8	98.5	4221 14	Ady17437 DNA encod
17	2789.8	98.5	4221 14	Ady61863 Human gen
18	2442.4	86.3	4083 12	Adq97696 Human Can

19	2290.6	80.9	3685	4	AAH17860	Aah17860 Human CDN
20	1833	64.7	1833	4	AAH14095	Aah14095 Human CDN
21	1520	53.7	1871	5	ABA82997	Aba82997 Human tra
22	1302.2	46.0	2426	12	ADQ97693	Adq97693 Mouse can
23	1273.8	45.0	1291	4	AAH14025	Aah14025 Human CDN
24	1032.2	36.5	1111	5	ADM19613	Adm19613 Novel hum
25	856.4	30.3	905	4	AAH05270	Aah05270 Human CDN
26	835.8	29.5	8892	12	ADQ97695	Adq97695 Human can
27	804.4	28.4	853	4	AAH05411	Aah05411 Human CDN
28	799.2	28.2	1365	5	ADM19367	Adm19367 Novel hum
29	649.8	23.0	57860	12	ADQ97692	Adq97692 Mouse can
30	613	21.7	745	12	ADQ97698	Adq97698 Human can
31	558	19.7	708	2	AAH98989	Aah98989 Human val
32	542.8	19.2	760	4	AAH06624	Aah06624 Human CDN
33	541	19.1	541	10	ADF80608	Adf80608 Leukaemia
34	518.6	18.3	593	4	AAH10633	Aah10633 Human CDN
35	507	17.9	507	10	AAD59581	Aad59581 Human gen
36	500.4	17.7	815	2	AAZ15581	Aaz15581 Human gen
37	451	15.9	451	14	ACL58353	ACL58353 Human col
38	440	15.5	468	9	ACH32745	ACH32745 Human end
39	439	15.5	543	5	ADL42840	Adl42840 Human ova
40	430.6	15.2	508	5	ADL44193	Adl44193 Human ova
41	421.4	14.9	445	4	AAH99154	Aah99154 83PSG4 su
42	421.4	14.9	477	2	AAZ17642	Aaz17642 Human gen
43	411.6	14.5	434	5	ADL37809	Adl37809 Human ova
44	411.6	14.5	434	5	ADI72670	Adi72670 Human ova
45	403.4	14.2	493	4	AAH33425	Aah33425 Human col

ALIGNMENTS

RESULT 1

ADf69739

ID ADF69739 standard; DNA; 2831 BP.

XX

AC ADF69739;

XX

DT 12-FEB-2004 (first entry)

XX

DE Gene encoding human retinoic acid (RA) regulated protein, 8.31.

XX

KW Human; retinoic acid regulated gene; retinoic acid regulated protein; RA;

KW mitogen; morphogen; stem cell progenitor cell;

KW embryonic brain development; lung; liver; kidney;

KW cancer cell differentiation; Ushers Disease; cancer; cytostatic;

KW chromosome 1q32.1-32.2; Gene; ds.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT CDS 124..2316

FT /*tag= a

FT /product= "Retinoic acid (RA) regulated protein, 8.31"

XX

PN US2003162268-A1.

 XX |

PD 28-AUG-2003.

 XX |

PF 08-APR-2003; 2003US-00409511.

 XX |

PR 14-JUL-1999; 99US-00354359.

 XX |

PA (IPNY/) IP N Y.

(CHEU/) W CHEUNG W M.

 XX |

PI Ip NY, Cheung WMM;

 XX |

DR WPI; 2003-897935/82.

DR P-PSDB; ADF69740.

 XX |

PT New retinoic acid regulated gene, useful in preparing a composition for treating or preventing Ushers Disease or cancer.

XX Claim 1; SEQ ID NO 1; 17pp; English.

XX The present invention relates to the isolation of a novel human retinoic acid (RA) regulated gene designated 8.31, and the polypeptide sequence it encodes. The gene encoding the 8.31 protein maps to human chromosome 1q32.1-32.2. The 8.31 protein is a mitogen or morphogen. It affects the identity of stem cell progenitor cells or the development of embryonic brain, lung, liver or kidney tissues. It also affects cancer cell differentiation. The 8.31 protein is useful in treating or preventing Ushers Disease or cancer. The present sequence encodes human retinoic acid (RA) regulated protein, 8.31.

XX Sequence 2831 BP; 807 A; 660 C; 607 G; 757 T; 0 U; 0 Other;

Query Match 100.0%; Score 2831; DB 10; Length 2831;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2831; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACGAGGGAGTTGGAGGCGATAACGATTTGTGTTGTGAGAGCGCAACGTGCGATT 60
DB |||||
QY 1 GGCACGAGGGAGTTGGAGGCGATAACGATTTGTGTTGTGAGAGCGCAACGTGCGATT 60
DB |||||

QY 61 TCTGCTGAACCTTGGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACC 120
DB |||||

QY 61 TCTGCTGAACCTTGGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACC 120
DB |||||

QY 121 CTGATGCTCTTCAATTCGGTGTCTCGCCAGCCAGCTTGGCGTCTGAGAAATGGATGG 180
DB |||||

QY 121 CTGATGCTCTTCAATTCGGTGTCTCGCCAGCCAGCTTGGCGTCTGAGAAATGGATGG 180
DB |||||

QY 181 TCTTCACAATACCCCTTCAATCCCTCTGACTGTTATCAGTGAGTGAGTGAATGATGAA 240
DB |||||

QY 181 TCTTCACAATACCCCTTCAATCCCTCTGACTGTTATCAGTGAGTGAGTGAATGATGAA 240
DB |||||

QY 241 CACACTTCTTATGGAGAACAGAGTCCAGTTCCTCTCTTGGATGACCTTCTCTTCT 300
DB |||||

QY 241 CACACTTCTTATGGAGAACAGAGTCCAGTTCCTCTCTTGGATGACCTTCTCTTCT 300
DB |||||

QY 301 GCTCCCAATATGGAACATGTACTAGCAGTTGCCAATGAAGAGGCTTTGTTGATTTGAT 360
DB |||||

QY 301 GCTCCCAATATGGAACATGTACTAGCAGTTGCCAATGAAGAGGCTTTGTTGATTTGAT 360
DB |||||

QY 361 AACACAGAAATCAAAAGTTTCAAAAGAGTGCTTCAAAAGATGAGTGAGTCACTGGAAT 420
DB |||||

QY 361 AACACAGAAATCAAAAGTTTCAAAAGAGTGCTTCAAAAGATGAGTGAGTCACTGGAAT 420
DB |||||

QY 421 GCGCTCTTGTGACCTGCGCTGGGTTCTGTGAACCTTAACTTTGTACAGCAGCGTGAT 480
DB |||||

QY 421 GCGCTCTTGTGACCTGCGCTGGGTTCTGTGAACCTTAACTTTGTACAGCAGCGTGAT 480
DB |||||

QY 481 CAAACAGCAAAATTTTGGGACGTAAAGAGTGGTGAGCTGATTGGAACATGCAAAAGTGAT 540
DB |||||

QY 481 CAAACAGCAAAATTTTGGGACGTAAAGAGTGGTGAGCTGATTGGAACATGCAAAAGTGAT 540
DB |||||

QY 541 CAATGAGCCTCAAGTCAAGTTCCTTTCTAAGTTTGAAGAGTGATTTCTGTACGGGT 600
DB |||||

QY 541 CAATGAGCCTCAAGTCAAGTTCCTTTCTAAGTTTGAAGAGTGATTTCTGTACGGGT 600
DB |||||

QY 601 GGAAGAGATGGCAACATTTATGCTGGGATACCGAGTGCAACAAAAGATGGGTTTTAT 660
DB |||||

QY 601 GGAAGAGATGGCAACATTTATGCTGGGATACCGAGTGCAACAAAAGATGGGTTTTAT 660
DB |||||

QY 661 AGGCAAGTGAATCAAAATCAGTGGAGCTCACATACCTCAGACAAAGCAACCCCTTCAAAA 720
DB |||||

QY 661 AGGCAAGTGAATCAAAATCAGTGGAGCTCACATACCTCAGACAAAGCAACCCCTTCAAAA 720
DB |||||

QY 721 CCCAAGAGAAACAGAAATTCAAAAGGACTTGTGCTCTCTGTGGATTTCCAGCAAAAGTGT 780
DB |||||

QY 721 CCCAAGAGAAACAGAAATTCAAAAGGACTTGTGCTCTCTGTGGATTTCCAGCAAAAGTGT 780
DB |||||

QY 781 ACTGTGGTCTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGATGGGATA 840
DB |||||

DB 781 ACTGTGGTCTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATA 840
QY 841 ATCAAAGTATGGGATTTACGTAAGAAATTAATCTGCTTATCGACAGAACCCATAGCATCC 900
DB |||||

DB 841 ATCAAAGTATGGGATTTACGTAAGAAATTAATCTGCTTATCGACAGAACCCATAGCATCC 900
QY 901 AAGTCTTTCTGTACCCAGGTAGCAGCACATCGAAACCTTCGATATTTCAAGTCTGATTTTG 960
DB |||||

DB 901 AAGTCTTTCTGTACCCAGGTAGCAGCACATCGAAACCTTCGATATTTCAAGTCTGATTTTG 960
QY 961 GATTCCACTGGCTCTACTTTTATTTGCTAAATTCACAGACGATACATCTACATGTTTAA 1020
DB |||||

DB 961 GATTCCACTGGCTCTACTTTTATTTGCTAAATTCACAGACGATACATCTACATGTTTAA 1020
QY 1021 ATGACTGGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACACCAAGAACTCTACCTTT 1080
DB |||||

DB 1021 ATGACTGGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACACCAAGAACTCTACCTTT 1080
QY 1081 TATGTAATAATCCAGCCTTAGTCCAGATGACCAAGTTTATTTAGTCTGCTCAAGTGAATGAA 1140
DB |||||

DB 1081 TATGTAATAATCCAGCCTTAGTCCAGATGACCAAGTTTATTTAGTCTGCTCAAGTGAATGAA 1140
QY 1141 GCTGCTTACATATGGAAGGCTCTCCACACCTCTGCAACCTCTCTACTGTGCTCTGGGTCT 1200
DB |||||

DB 1141 GCTGCTTACATATGGAAGGCTCTCCACACCTCTGCAACCTCTCTACTGTGCTCTGGGTCT 1200
QY 1201 TCTCAAAGGTCAGCTCTGTGTGCTGTGTGCTCATCTGACTTCAAAAGATTTGCTACCTGT 1260
DB |||||

DB 1201 TCTCAAAGGTCAGCTCTGTGTGCTGTGTGCTCATCTGACTTCAAAAGATTTGCTACCTGT 1260
QY 1261 TCTGATGACAAATACACTAAAAAATCTGGCGTTGTAATAGAGCTTTAGAGGAGAAACACGGA 1320
DB |||||

DB 1261 TCTGATGACAAATACACTAAAAAATCTGGCGTTGTAATAGAGCTTTAGAGGAGAAACACGGA 1320
QY 1321 GGTGATAAACTTTCCACGGTGGGTGGGCTCTCAGAAAGAAAAAGAGTCAAGACCTGGC 1380
DB |||||

DB 1321 GGTGATAAACTTTCCACGGTGGGTGGGCTCTCAGAAAGAAAAAGAGTCAAGACCTGGC 1380
QY 1381 CTAGTAAACAGTAGCAGTAGCAGTAGTCTCTGCGCCAAAGCCCCCAGGGTAAAGTGAAT 1440
DB |||||

DB 1381 CTAGTAAACAGTAGCAGTAGCAGTAGTCTCTGCGCCAAAGCCCCCAGGGTAAAGTGAAT 1440
QY 1441 CCATCCAAATTTCTCCCGTCTATCCGAGCTTGTGCCCCCAAGCTGTGCTGGAGACCTCCCT 1500
DB |||||

DB 1441 CCATCCAAATTTCTCCCGTCTATCCGAGCTTGTGCCCCCAAGCTGTGCTGGAGACCTCCCT 1500
QY 1501 CTTCTTTCAAAATCTCCTACGTTCTCTATTAAACCTCTCTGCAAGGCCCCGGTCTCCC 1560
DB |||||

DB 1501 CTTCTTTCAAAATCTCCTACGTTCTCTATTAAACCTCTCTGCAAGGCCCCGGTCTCCC 1560
QY 1561 ATCAACAGAGAGGCTCTGTCTCTCCGTCTCTCCAGAGCACCTTCTCATCTTTCAAGATG 1620
DB |||||

DB 1561 ATCAACAGAGAGGCTCTGTCTCTCCGTCTCTCCAGAGCACCTTCTCATCTTTCAAGATG 1620
QY 1621 TCGATTAGAACTGGGTGACCCGAAACCTTCTCATCACCCACCTCATCTCCACCTGCT 1680
DB |||||

DB 1621 TCGATTAGAACTGGGTGACCCGAAACCTTCTCATCACCCACCTCATCTCCACCTGCT 1680
QY 1681 TCGGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCAATCC 1740
DB |||||

DB 1681 TCGGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCAATCC 1740
QY 1741 CAAGCAGAGGCTTCTCTGAGTCTAGAAATAGAGTAAAGAGGAGCTAGACTCAAGCTGT 1800
DB |||||

DB 1741 CAAGCAGAGGCTTCTCTGAGTCTAGAAATAGAGTAAAGAGGAGCTAGACTCAAGCTGT 1800
QY 1801 CTGAGAGTGTGAACAAAAGTGTGAAGAGTTGTAACCTGTGTGACTGAGCTTTGATGGC 1860
DB |||||

DB 1801 CTGAGAGTGTGAACAAAAGTGTGAAGAGTTGTAACCTGTGTGACTGAGCTTTGATGGC 1860
QY 1861 CAAGTTGAAATCTTCAATTTGGATCTGTGCTGCTTGTGTAACCCAGGAGACCTTAGT 1920
DB |||||

DB 1861 CAAGTTGAAATCTTCAATTTGGATCTGTGCTGCTTGTGTAACCCAGGAGACCTTAGT 1920

```
QY 1921 AAGGACTCTTAGTCTCTACCAAACTCAAGCAAAATTTGAAGGAGCTGGTACCAAGTATCTCA 1980
Db 1921 AAGGACTCTTAGTCTCTACCAAACTCAAGCAAAATTTGAAGGAGCTGGTACCAAGTATCTCA 1980
QY 1981 GAGCCTCCGCTCTCTATCAGTCCGATGCTTGCAGAAAGCTGTGGAACGCTACCTCTTCT 2040
Db 1981 GAGCCTCCGCTCTCTATCAGTCCGATGCTTGCAGAAAGCTGTGGAACGCTACCTCTTCT 2040
QY 2041 TTGAGACCTTTGGAGAGGGTCTGAAATGGTAGGCAAGAGAAATAGTTCCCAAGAAAT 2100
Db 2041 TTGAGACCTTTGGAGAGGGTCTGAAATGGTAGGCAAGAGAAATAGTTCCCAAGAAAT 2100
QY 2101 AAAAACTGGTGTGGCCATGCGAGCCAAACGGAAGGCTGAGAATCCATCTCCAGAACT 2160
Db 2101 AAAAACTGGTGTGGCCATGCGAGCCAAACGGAAGGCTGAGAATCCATCTCCAGAACT 2160
QY 2161 CCGTCATCCAGACACCCCAATTCAGGAGACAGAGCGGAAGACATTCGCAAGCCGGTTC 2220
Db 2161 CCGTCATCCAGACACCCCAATTCAGGAGACAGAGCGGAAGACATTCGCAAGCCGGTTC 2220
QY 2221 ACCATCAGCCAGCTCCATGAGGAAATCTGCACATACCTTCATAGAAAGTCCCGAGAG 2280
Db 2221 ACCATCAGCCAGCTCCATGAGGAAATCTGCACATACCTTCATAGAAAGTCCCGAGAG 2280
QY 2281 GACTTCTGTGGTCTGGAACACTCAACAGAAATATAGATTTCTAATCTGAGTGAAGTACTCA 2340
Db 2281 GACTTCTGTGGTCTGGAACACTCAACAGAAATATAGATTTCTAATCTGAGTGAAGTACTCA 2340
QY 2341 GCTTTGGTCCACTAAAAAAGCTGAGCTTTGGTCCACTAAAAAAGATGAAAAATACAAAG 2400
Db 2341 GCTTTGGTCCACTAAAAAAGCTGAGCTTTGGTCCACTAAAAAAGATGAAAAATACAAAG 2400
QY 2401 AGTGACTCTAATCTGTGCTTTTAAGAAAGCTGCTTTTCAATTTTGAACAAAATCTTTT 2460
Db 2401 AGTGACTCTAATCTGTGCTTTTAAGAAAGCTGCTTTTCAATTTTGAACAAAATCTTTT 2460
QY 2461 TCAAGCTCAAAATGACCTAATCTGCTTCTACTACCATATGATATGATGAGCTTCCCGAG 2520
Db 2461 TCAAGCTCAAAATGACCTAATCTGCTTCTACTACCATATGATATGATGAGCTTCCCGAG 2520
QY 2521 GATGAATGCTGTGCTTTTAAATTTTCAAAAGTAAATTTGTCACTCTAGCATTTTGAATGAAT 2580
Db 2521 GATGAATGCTGTGCTTTTAAATTTTCAAAAGTAAATTTGTCACTCTAGCATTTTGAATGAAT 2580
QY 2581 AGCTTTCACCTTTTAAATTTTCAATCTCTCTATATAATGACATCCCAAGTTTCATGGAGG 2640
Db 2581 AGCTTTCACCTTTTAAATTTTCAATCTCTCTATATAATGACATCCCAAGTTTCATGGAGG 2640
QY 2641 CAAAAAACAAGTTTCTTGTATCTCTGAAACTTTCTATGCTCAGTGAAGATATCTGCCAG 2700
Db 2641 CAAAAAACAAGTTTCTTGTATCTCTGAAACTTTCTATGCTCAGTGAAGATATCTGCCAG 2700
QY 2701 CCACAGCATGAGGCTGTGAAGCTGACCTGAGAAATCTCTGCTGAAGACCCCTGGTCTCT 2760
Db 2701 CCACAGCATGAGGCTGTGAAGCTGACCTGAGAAATCTCTGCTGAAGACCCCTGGTCTCT 2760
QY 2761 GTTCTGCTCCCAACATGATAATTTTATTTGAAATAATAATCTTTTCACTATGAAAAAA 2820
Db 2761 GTTCTGCTCCCAACATGATAATTTTATTTGAAATAATAATCTTTTCACTATGAAAAAA 2820
QY 2821 AAAAAAATAA 2831
Db 2821 AAAAAAATAA 2831
```

RESULT 2

ADX69660

ID ADX69660 standard; cDNA; 2831 BP.

XX

AC

XX

DT

05-MAY-2005 (first entry)

```
XX Human retinoic acid regulated nuclear matrix protein (RAMP) cDNA.
DE DNA purification; retinoic acid; diagnosis; hepatocellular carcinoma;
KW cytosolic; prognosis; cancer; liver tumor;
KW retinoic acid regulated nuclear matrix protein; gene; ds.
XX Homo sapiens.
XX Key Location/Qualifiers
FT 124..2316
FT /*tag= a
FT /product= "retinoic acid regulated nuclear matrix
FT protein"
XX US2005037372-A1.
XX 17-FEB-2005.
XX 02-DEC-2003; 2003US-00726160.
XX 14-JUL-1999; 99US-00354359.
XX 08-APR-2003; 2003US-00409511.
XX (UHK-) UNIV HONG KONG SCI & TECHNOLOGY.
XX Ip NY, Cheung WNW;
XX WPI; 2005-172246/18.
XX P-PSDB; ADX69661.
XX Novel isolated retinoic acid regulated gene, useful for screening and
XX determining prognosis of patient having Hepatocellular cancer.
XX Claim 1; SEQ ID NO 1; 21bp; English.
XX The invention relates to an isolated retinoic acid regulated gene (I)
XX having a fully defined 2831 nucleotides sequence given in specification.
XX A protein (II) encoded by (I) is useful as a screening tool for
XX diagnosing Hepatocellular carcinomas and for monitoring treatment or
XX progression of Hepatocellular carcinomas. An antibody (III) to (II) is
XX useful for screening and determining the prognosis of a patient having
XX Hepatocellular cancer, where the biological samples comprise liver
XX tissues. This sequence corresponds to the cDNA encoding the retinoic acid
XX regulated nuclear matrix protein (RAMP; I).
XX Sequence 2831 BP; 807 A; 660 C; 607 G; 757 T; 0 U; 0 Other;
QY Query Match 100.0%; Score 2831; DB 14; Length 2831;
QY Best Local Similarity 100.0%; Pred. No. 0;
QY Matches 2831; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCACGACGGGAGTTGGAGCGGATACGATTTGTTGTGTGAGAGCGCAACGTCGGATT 60
Db 1 GGCACGACGGGAGTTGGAGCGGATACGATTTGTTGTGTGAGAGCGCAACGTCGGATT 60
QY 61 TCTGCTGAACCTTGAGGCGATTTCTACGACTTTCTCTCAGCTGAGGCTTTCTCCGACC 120
Db 61 TCTGCTGAACCTTGAGGCGATTTCTACGACTTTCTCTCAGCTGAGGCTTTCTCCGACC 120
QY 121 CTGATGCTCTTCAATTCGGTGTCTCCGACGCCAGCTTGGCGTCTCAGAAATGGATGG 180
Db 121 CTGATGCTCTTCAATTCGGTGTCTCCGACGCCAGCTTGGCGTCTCAGAAATGGATGG 180
QY 181 TCTTCACAATACCTCTTCAATCCCTTCTGACTGGTATCAGTGCAGTGGTAAATGATGA 240
Db 181 TCTTCACAATACCTCTTCAATCCCTTCTGACTGGTATCAGTGCAGTGGTAAATGATGA 240
QY 241 CACACTTCTTATGAGAAACAGGAGTCCAGTTCCTCTCTTTGGATGATACCTTCTCTCT 300
Db 241 CACACTTCTTATGAGAAACAGGAGTCCAGTTCCTCTCTTTGGATGATACCTTCTCTCT 300
QY 301 GCTCCCAATATGGAACATGTACTAGCAGTTGCCAATGAAGAGGCTTTGTTGCGATTGAT 360
```

||||| 301 GCTCCCAATATGGAACATGTACTAGAGTTGCCAATGAAGAGCTTTGTTGATGAT 360
QY 361 AACACAGAAATCAAAAGTTTCAGAAAGAGTGCTTTCAAGAAATGATGGCTCACTGGAA 420
Db 361 AACACAGAAATCAAAAGTTTCAGAAAGAGTGCTTTCAAGAAATGATGGCTCACTGGAA 420
QY 421 GCGGCTTTGACCTGGCCCTGGGTTCTGGTGAACCTTAAACTTTGTTACAGCAGCAGGTGAT 480
Db 421 GCGGCTTTGACCTGGCCCTGGGTTCTGGTGAACCTTAAACTTTGTTACAGCAGCAGGTGAT 480
QY 481 CAAACAGCCAAATTTTGGGACGTAAAGCTGGTGAGCTGATTGGAAACATGCAAAAGGTGAT 540
Db 481 CAAACAGCCAAATTTTGGGACGTAAAGCTGGTGAGCTGATTGGAAACATGCAAAAGGTGAT 540
QY 541 CAATCAGCCTCAAGTCAAGTTGCGCTTTTCTAAGTTTGAGAAAGCTGTATTTCTGTACGGGT 600
Db 541 CAATCAGCCTCAAGTCAAGTTGCGCTTTTCTAAGTTTGAGAAAGCTGTATTTCTGTACGGGT 600
QY 601 GGAAGAGATGGCAACATTTATGTCTGGGATACCAGGTGCAACAAAAGATGGGTTTTAT 660
Db 601 GGAAGAGATGGCAACATTTATGTCTGGGATACCAGGTGCAACAAAAGATGGGTTTTAT 660
QY 661 AGGCAAGTGAATCAAAATCAGTGGAGCTCAACAATACCTCAGACAAGCAAAACCCCTTCAAAA 720
Db 661 AGGCAAGTGAATCAAAATCAGTGGAGCTCAACAATACCTCAGACAAGCAAAACCCCTTCAAAA 720
QY 721 CCCAAGAGAAACAGAAATTCAGAAAGACCTGCTCTCTGTGGATTTCCAGCAAAAGTGT 780
Db 721 CCCAAGAGAAACAGAAATTCAGAAAGACCTGCTCTCTGTGGATTTCCAGCAAAAGTGT 780
QY 781 ACTGTGGTCTCTTTCAGACAGAGATACCTTTAGTCTCAGCAGGAGCTGTGGATGGGATA 840
Db 781 ACTGTGGTCTCTTTCAGACAGAGATACCTTTAGTCTCAGCAGGAGCTGTGGATGGGATA 840
QY 841 ATCAAAGTATGGATTTACGTAAAGATATATCTGTTATCGACAAGAACCCATGATCATCC 900
Db 841 ATCAAAGTATGGATTTACGTAAAGATATATCTGTTATCGACAAGAACCCATGATCATCC 900
QY 901 AAGTCTTTCTGTATCCAGGTAGCAGCACTCGAAACCTTGGATATTCAAAGTCTGATTTTG 960
Db 901 AAGTCTTTCTGTATCCAGGTAGCAGCACTCGAAACCTTGGATATTCAAAGTCTGATTTTG 960
QY 961 GATTCACCTGGCTCTACTTTATTTGCTAATGTCACAGACGATAACATCTACATGTTTAA 1020
Db 961 GATTCACCTGGCTCTACTTTATTTGCTAATGTCACAGACGATAACATCTACATGTTTAA 1020
QY 1021 ATGACTGGGTTCAAGACTTCTCCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTT 1080
Db 1021 ATGACTGGGTTCAAGACTTCTCCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTT 1080
QY 1081 TATGTAAATCCAGCTTTAGTCCAGATGACCAAGTTTATGTCAGTGGCTCAAGTGAATGAA 1140
Db 1081 TATGTAAATCCAGCTTTAGTCCAGATGACCAAGTTTATGTCAGTGGCTCAAGTGAATGAA 1140
QY 1141 GCTGCTACATATGAAGGTCTCCACACCTGGGAAACCTCTACTGTGCTCTGGGTGAT 1200
Db 1141 GCTGCTACATATGAAGGTCTCCACACCTGGGAAACCTCTACTGTGCTCTGGGTGAT 1200
QY 1201 TCTCAAGAGGTCAAGTCTGTGTGCTGGTGTCCATCTGACTTCCAAAGATTTGCTACCTGT 1260
Db 1201 TCTCAAGAGGTCAAGTCTGTGTGCTGGTGTCCATCTGACTTCCAAAGATTTGCTACCTGT 1260
QY 1261 TCTGATGACAAATACATACTAAAACTGGGCTTGAATAGAGGCTTAGAGAGAGAAACAGGA 1320
Db 1261 TCTGATGACAAATACATACTAAAACTGGGCTTGAATAGAGGCTTAGAGAGAGAAACAGGA 1320
QY 1321 GGTGATAACTTTTCCAGGTGGGTTGGGCTCTCAGAGAGAAAGATCAAGACTTGGC 1380
Db 1321 GGTGATAACTTTTCCAGGTGGGTTGGGCTCTCAGAGAGAAAGATCAAGACTTGGC 1380
QY 1381 CTAGTAAACAGTAAACAGTACAGAGTACTCTGCTGCCAAAGCCCCAGGGTAAAGTGAAT 1440
|||||

Db 1381 CTAGTAAACAGTAAACAGTAGCCAGAGTACTCTCTGCCAAAGCCCCAGGGTAAAGTGAAT 1440
QY 1441 CCATCCAAATTTTCCCGTCTATCCGJAGCTTGTGCCCCAAGCTGTGTGGAGACTTCCCT 1500
Db 1441 CCATCCAAATTTTCCCGTCTATCCGJAGCTTGTGCCCCAAGCTGTGTGGAGACTTCCCT 1500
QY 1501 CTTCTTTCAAAATCTCTCTAGTCTCTATTTAAAACTCTCTCTGCAAGGCCCGGTCTCCC 1560
Db 1501 CTTCTTTCAAAATCTCTCTAGTCTCTATTTAAAACTCTCTCTGCAAGGCCCGGTCTCCC 1560
QY 1561 ATCAACAGAGAGGCTCTCTCTCTCTCTCTCCCAAGCCACCTTCACTTTTCAAGATG 1620
Db 1561 ATCAACAGAGAGGCTCTCTCTCTCTCTCTCTCTCTCTCCCAAGCCACCTTCACTTTTCAAGATG 1620
QY 1621 TCGATTAGAAACTCGGGTGACCCGAAACACCTTCTCTCATCACCCATCACTCCGACTGCT 1680
Db 1621 TCGATTAGAAACTCGGGTGACCCGAAACACCTTCTCTCATCACCCATCACTCCGACTGCT 1680
QY 1681 TCGGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCACTCC 1740
Db 1681 TCGGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCACTCC 1740
QY 1741 CAAGCAGAGGCTTCTCTCTGAGTCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGT 1800
Db 1741 CAAGCAGAGGCTTCTCTCTGAGTCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGT 1800
QY 1801 CTGAGAGTGTGAAACAAAAGTGTGAGAGTTGTAACCTGTGTGACTGAGCTTGTATGGC 1860
Db 1801 CTGAGAGTGTGAAACAAAAGTGTGAGAGTTGTAACCTGTGTGACTGAGCTTGTATGGC 1860
QY 1861 CAAGTGTGAATCTTCAATTTGGATCTGTCTGCTGCTCTGCTGTAACACAGGAAGACTTGT 1920
Db 1861 CAAGTGTGAATCTTCAATTTGGATCTGTCTGCTGCTCTGCTGTAACACAGGAAGACTTGT 1920
QY 1921 AAGGACTCTCTAGTCTCTTACCAAAATCAAGCAAAATTTGAAGGAGCTGGTACCAGTATCTCA 1980
Db 1921 AAGGACTCTCTAGTCTCTTACCAAAATCAAGCAAAATTTGAAGGAGCTGGTACCAGTATCTCA 1980
QY 1981 GAGCTCTGCTCTCTATCAGTCTGTTTCAAGAAAGCTGTGGAAACGCTACCTCTTCTCT 2040
Db 1981 GAGCTCTGCTCTCTATCAGTCTGTTTCAAGAAAGCTGTGGAAACGCTACCTCTTCTCT 2040
QY 2041 TTGAGACTCTGTGAGAGAGGCTCTGAAATGTTAGGCAAGAGAAATAGTTCCCCAGAGAA 2100
Db 2041 TTGAGACTCTGTGAGAGAGGCTCTGAAATGTTAGGCAAGAGAAATAGTTCCCCAGAGAA 2100
QY 2101 AAAAATCTGTTGTTGGCCATGGCAGCCAAACGGAAGGCTGAGAATCCATCTCCACGAAGT 2160
Db 2101 AAAAATCTGTTGTTGGCCATGGCAGCCAAACGGAAGGCTGAGAATCCATCTCCACGAAGT 2160
QY 2161 CCGTCTATCCAGACACCCAAATTCAGGAGACAGAGCGGAAAGACTTGGCCAGGCCGGTC 2220
Db 2161 CCGTCTATCCAGACACCCAAATTCAGGAGACAGAGCGGAAAGACTTGGCCAGGCCGGTC 2220
QY 2221 ACCATCAGCCCGCTCCATGAGGAAATCTGCACATCTTCCATAGAAAGTCCAGGAG 2280
Db 2221 ACCATCAGCCCGCTCCATGAGGAAATCTGCACATCTTCCATAGAAAGTCCAGGAG 2280
QY 2281 GACTTCTGTGCTCTGAACTCAACAGACTCAACAGAAATTAAGATTTCTAACTGAGTGAATCT 2340
Db 2281 GACTTCTGTGCTCTGAACTCAACAGACTCAACAGAAATTAAGATTTCTAACTGAGTGAATCT 2340
QY 2341 GCTTTGGTCCATTAACCAAGCTGAGCTTTGGTCCACTTAAACAGAGTGAAGAAATACAG 2400
Db 2341 GCTTTGGTCCATTAACCAAGCTGAGCTTTGGTCCACTTAAACAGAGTGAAGAAATACAG 2400
QY 2401 AGTGACTCTATAACTCTGGTCTTTAAGAAAGTGGCTTTTCACTTTTAGACAAAATCTTT 2460
Db 2401 AGTGACTCTATAACTCTGGTCTTTAAGAAAGTGGCTTTTCACTTTTAGACAAAATCTTT 2460
QY 2461 TCAACGCTGAAATGATCCTAACTCTGTTTCTACTACCAATATGATATGACAGCTTCCCGAG 2520
Db 2461 TCAACGCTGAAATGATCCTAACTCTGTTTCTACTACCAATATGATATGACAGCTTCCCGAG 2520

QY 2521 GATGAATGCTGTGTTTAAATTTTAAAGTAAATTTGTCTACCTAGCATTTTGAATGAAT 2580
DB 2521 GATGAATGCTGTGTTTAAATTTTAAAGTAAATTTGTCTACCTAGCATTTTGAATGAAT 2580
QY 2581 AGTCTTCACTTTTAAATTTATTCATCTTCTATATAATGACATCCCGAGTTTCATGGAGG 2640
DB 2581 AGTCTTCACTTTTAAATTTATTCATCTTCTATATAATGACATCCCGAGTTTCATGGAGG 2640
QY 2641 CAAAAACAAGTTCTTGTTATCTCTGAAACTTCTATGCTCAGTGAAGATATCTGCCAG 2700
DB 2641 CAAAAACAAGTTCTTGTTATCTCTGAAACTTCTATGCTCAGTGAAGATATCTGCCAG 2700
QY 2701 CCACAGCATGAGGCTGTGAAGCTGACCTGAGAAATCCTCTGCTGAAGACCCCTGTTCT 2760
DB 2701 CCACAGCATGAGGCTGTGAAGCTGACCTGAGAAATCCTCTGCTGAAGACCCCTGTTCT 2760
QY 2761 GTTCTGCCCTCCACATGATATAATTTTATTTGAAATAACATAATCTTTCTACTATGAAAAA 2820
DB 2761 GTTCTGCCCTCCACATGATATAATTTTATTTGAAATAACATAATCTTTCTACTATGAAAAA 2820
QY 2821 AAAAAA 2831
DB 2821 AAAAAA 2831

RESULT 3

AAH99153
ID AAH99153 standard; cDNA; 2838 BP.

AC AAH99153;

XX 04-DEC-2001 (first entry)

DE Human prostate-related gene 83P5G4 cDNA.

XX 83P5G4; PCR primer; DNA adaptor; prostate; testis; tissue; cancer; ss;
KW tumour; kidney; brain; bone; ovary; breast; pancreas; uterus; colon;
KW lung; cytostatic; gene therapy; antibody therapy; ribozyme; liver;
KW single chain monoclonal antibody; serum; blood; urine; bladder; cervix;
KW rectum; stomach; human; chromosome 1q31-q32.

XX Homo sapiens.

XX WO200159115-A2.

XX 16-AUG-2001.

XX 09-FEB-2001; 2001WO-US004426.

XX 09-FEB-2000; 2000US-0181261P.

XX (UROC-) UROGENESYS INC.

XX Hubert RS, Afar DEH, Challita-Eid PM, Faris M, Levin B;

PI Mitchell SC, Jakobovits A;

XX WPI; 2001-514669/56.

DR P-PSDB; AAM25224.

XX An isolated 83P5G4-related protein useful as a diagnostic and/or
PT therapeutic agent in multiple cancers such as prostate, bladder and bone
PT cancer.

XX Claim 1; Fig 2; 112pp; English.

XX The nucleic acid sequences represent the 83P5G4 gene and the primers and
CC adaptors used to amplify 83P5G4 DNA. 83P5G4 exhibits prostate specific
CC expression in normal adult tissue, but it is also aberrantly expressed in
CC many cancers including tumours of the prostate, testis, bladder, kidney,
CC brain, bone, cervix, uterus, ovary, breast, pancreas, stomach, rectum,
CC liver, colon and lung. The 83P5G4 polynucleotide, its related protein and
CC also peptide fragments of the protein are therefore useful for diagnosing

CC and treating cancer. A vector comprising a polynucleotide which encodes a
CC single chain monoclonal antibody, that immunospecifically binds to an
CC 83P5G4-related protein, and a ribozyme capable of cleaving a
CC polynucleotide having the 83P5G4 coding sequence, are both useful in the
CC preparation of a composition for treating a patient with a cancer that
CC expresses 83P5G4. The sequences can be used in diagnostic methods to
CC monitor the level of 83P5G4 gene products in serum, blood, urine and
CC tissue and to thereby detect the presence of cancerous cells

XX Sequence 2838 BP; 810 A; 661 C; 609 G; 758 T; 0 U; 0 Other;

QY Query Match 99.5%; Score 2816.2; DB 4; Length 2838;

DB Best Local Similarity 99.7%; Pred. No. 0;

XX Matches 2821; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 3 CACGAGCGGAGTTGGAGCGGATTAACGATTTGTGTGTGAGAGGCGCAACGTCCGATTTTC 62

DB 9 CAGTGGCGGAGTTGGAGCGGATTAACGATTTGTGTGTGAGAGGCGCAACGTCCGATTTTC 68

QY 63 TGCTGAATCTTGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCT 122

DB 69 TGCTGAATCTTGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCT 128

QY 123 GATGCTCTTCAATTCGGTGTCTCCGACGCCAGCTTGGCGTCTCTGAGAAATGGATGTC 182

DB 129 GATGCTCTTCAATTCGGCGCTCCGCGCAGCTTGGCGTCTCTGAGAAATGGATGTC 188

QY 183 TTCACAATACCTCTTCAATCCCTTCTGACTGGTTATCAGTCAGAGTGTATGATGAACA 242

DB 189 TTCACAATACCTCTTCAATCCCTTCTGACTGGTTATCAGTCAGTGTATGATGAACA 248

QY 243 CACTTTCTTATGAGAAACAGAGTCCGAGTTCCTCTTTTGGATGATCTTCTCTCTTCGC 302

DB 249 CACTTTCTTATGAGAAACAGAGTCCGAGTTCCTCTTTTGGATGATCTTCTCTCTTCGC 308

QY 303 TCCCAATATGGAACATGTACTAGCAGTTGCCAATGAAGAAGCTTTGTTGATTTGATTA 362

DB 309 TCCCAATATGGAACATGTACTAGCAGTTGCCAATGAAGAAGCTTTGTTGATTTGATTA 368

QY 363 CACAGAATCACAAAGTTTTCAGAAAGAGTGTCTTCAAGAAATGGATGGCTCATGGAATGC 422

DB 369 CACAGAATCACAAAGTTTTCAGAAAGAGTGTCTTCAAGAAATGGATGGCTCATGGAATGC 428

QY 423 CGTCTTTGACCTGGCGTCTGGTTCCTGTTGAACTTAACTTTTACAGCAGCAGGTGATCA 482

DB 429 CGTCTTTGACCTGGCGTCTGGTTCCTGTTGAACTTAACTTTTACAGCAGCAGGTGATCA 488

QY 483 AACAGCCAAATTTTGGGACGTAAAGAGCTGGTGGATTTGSAACATGCAAGGTTCATCA 542

DB 489 AACAGCCAAATTTTGGGACGTAAAGAGCTGGTGGATTTGSAACATGCAAGGTTCATCA 548

QY 543 ATGAGCCTCAAGTCAGTTGCTTTTCTAAGTTTGAAGAAAGCTGTATTTCTGACGGTGG 602

DB 549 ATGAGCCTCAAGTCAGTTGCTTTTCTAAGTTTGAAGAAAGCTGTATTTCTGACGGTGG 608

QY 603 AAGAGATGCGCAACATTATGCTCTGGATACAGAGTGCACAAAGAAAGATGGGTTTATAG 662

DB 609 AAGAGATGCGCAACATTATGCTCTGGATACAGAGTGCACAAAGAAAGATGGGTTTATAG 668

QY 663 GCAAGTGAATCAAATCAGTGGAGCTCACAATACTCTCAGACAGCAAAACCCCTTCAAAACC 722

DB 669 GCAAGTGAATCAAATCAGTGGAGCTCACAATACTCTCAGACAGCAAAACCCCTTCAAAACC 728

QY 723 CAAGAAGAAACAGAAATTCAAAAGGACTTGTCTCTGTTGAAATTTCCAGCAAGGTGTAC 782

DB 729 CAAGAAGAAACAGAAATTCAAAAGGACTTGTCTCTGTTGAAATTTCCAGCAAGGTGTAC 788

QY 783 TGTGCTCTCTTCAAGACAGAAATACCTTAGTCTCAGCAGAGCTGTGGATGGGATAT 842

DB 789 TGTGCTCTCTTCAAGACAGAAATACCTTAGTCTCAGCAGAGCTGTGGATGGGATAT 848

QY 843 CAAAGTATGGGATTTTACCGTAAGAATTTATCTGCTTATCGACAAGAACCCATAGCATCAA 902

849 CAAAGTATGGGATTACGTAAGAAATTATACGTCTTATCGACAAGAACCCATAGCATCCAA 908
903 GTCTTTCTCTGATACCCAGGTAGCAGACCTCGAAAACTTGGATATTCAAGTCTGATTTTGGGA 962
909 GTCTTTCTCTGATACCCAGGTAGCAGACCTCGAAAACTTGGATATTCAAGTCTGATTTTGGGA 968
963 TTCCCACTGGCTCTACTTTATTTGCTAAATTTGACACAGCATACATCTACATGTTTAAATAT 1022
969 TTCCCACTGGCTCTACTTTATTTGCTAAATTTGACACAGCATACATCTACATGTTTAAATAT 1028
1023 GACTGGGTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACACAGAACTCTACCTTTTAA 1082
1029 GACTGGGTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACACAGAACTCTACCTTTTAA 1088
1083 TGTAAAAATCCAGCCTTAGTCCAGATGACACAGTFTTTTAGTCAGTGGCTCAAGTGATGAAGC 1142
1089 TGTAAAAATCCAGCCTTAGTCCAGATGACACAGTFTTTTAGTCAGTGGCTCAAGTGATGAAGC 1148
1143 TGCCTACATATGGAGGTCTCCACACCTTGGCAACCTCTACTGTGCTCCTGGGTCAATTC 1202
1149 TGCCTACATATGGAGGTCTCCACACCTTGGCAACCTCTACTGTGCTCCTGGGTCAATTC 1208
1203 TCAAGAGGTCAAGTCTGTGTGCTGGTCCATCTGACTTTCACAAAGATTTGCTACCTGTTTC 1262
1209 TCAAGAGGTCAAGTCTGTGTGCTGGTCCATCTGACTTTCACAAAGATTTGCTACCTGTTTC 1268
1263 TGAATGACAAATACACTAAAAATCTGGCGTTGGAATAGAGGCTTAGAGGAGAAACCAAGGAGG 1322
1269 TGAATGACAAATACACTAAAAATCTGGCGTTGGAATAGAGGCTTAGAGGAGAAACCAAGGAGG 1328
1323 TGAATAAATCTTTCCAGGTGGTGGCTCTCAGAGAAAGAGTCAAGACCTGGCGCT 1382
1329 TGAATAAATCTTTCCAGGTGGTGGCTCTCAGAGAAAGAGTCAAGACCTGGCGCT 1388
1383 AGTAACAGTAAACAGGTAGCCAGAGTACTCTGCCAAAGCCCCAGGGTAAAGTGCAATCC 1442
1389 AGTAACAGTAAACAGGTAGCCAGAGTACTCTGCCAAAGCCCCAGGGTAAAGTGCAATCC 1448
1443 ATCCAAATCTTTCCCGTCAATCCGAGTGTGTCGCCCAAGCTGTGTGAGAGCTCCCTCT 1502
1449 ATCCAAATCTTTCCCGTCAATCCGAGTGTGTCGCCCAAGCTGTGTGAGAGCTCCCTCT 1508
1503 TCCTTCAATACCTCTAGTCTCTATTAAACCTCTCTGCCAAGGCCGCTCTCCCAT 1562
1509 TCCTTCAATACCTCTAGTCTCTATTAAACCTCTCTGCCAAGGCCGCTCTCCCAT 1568
1563 CAAAGAGAGGCTCTGTCTCTCCGTCTCTCCCAAGCACCTTCACTCTTTCAAGATGTC 1622
1569 CAAAGAGAGGCTCTGTCTCTCCGTCTCTCCCAAGCACCTTCACTCTTTCAAGATGTC 1628
1623 GATTAGAAATCTGGGTGACCCGAAACCTTCTCATCAACCCATCACTCCACCTGGCTTC 1682
1629 GATTAGAAATCTGGGTGACCCGAAACCTTCTCATCAACCCATCACTCCACCTGGCTTC 1688
1683 GGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATCTGTGAGCCAGAGTCAATCCCA 1742
1689 GGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATCTGTGAGCCAGAGTCAATCCCA 1748
1743 AGCAGAGGCTTGTCTGTAGTCTAGAAATAGAGTAAGAGGAGGCTAGACTCAAGCTGTCT 1802
1749 AGCAGAGGCTTGTCTGTAGTCTAGAAATAGAGTAAGAGGAGGCTAGACTCAAGCTGTCT 1808
1803 GGAGAGTGTGAACAAAAAGTGTGTGAAGTGTGAATCTGTGTGACTGAGCTTGATGGCCA 1862
1809 GGAGAGTGTGAACAAAAAGTGTGTGAAGTGTGAATCTGTGTGACTGAGCTTGATGGCCA 1868
1863 AGTTGAAATCTTCAATTTGGATCTGTGCTGGCTTCTGGTAACACAGGAAGCCCTTAGTAA 1922
1869 AGTTGAAATCTTCAATTTGGATCTGTGCTGGCTTCTGGTAACACAGGAAGCCCTTAGTAA 1928
1923 GGACTCTCTAGTCTCTACCAATCAAGCAAAATTTGAAGAGGCTGTGTACAGTATCTCAGA 1982
1929 GGACTCTCTAGTCTCTACCAATCAAGCAAAATTTGAAGAGGCTGTGTACAGTATCTCAGA 1988

1983 GCCTCCGTCTCTATCAGTCCGTATCTCAGAAAGCTGTGGAACGCTACCTCTCTCTTTT 2042
1989 GCCTCCGTCTCTATCAGTCCGTATCTCAGAAAGCTGTGGAACGCTACCTCTCTCTTTT 2048
2043 GAGACCTTTGGGAGAGGGTCTGAAATGGTAGGCAAGAGAAATAGTTCCCCAGAGAAATAA 2102
2049 GAGACCTTTGGGAGAGGGTCTGAAATGGTAGGCAAGAGAAATAGTTCCCCAGAGAAATAA 2108
2103 AAACCTGGTGTGTGGCCATGGCAGCCAAACGGAAGGCTGAGAATCCATCTCCACAAGTCC 2162
2109 AAACCTGGTGTGTGGCCATGGCAGCCAAACGGAAGGCTGAGAATCCATCTCCACAAGTCC 2168
2163 GTATCCAGACACCCAAATTTCCAGAGACAGAGGCGGAAGACATTTGCCAAGCCCCGCTCAC 2222
2169 GTATCCAGACACCCAAATTTCCAGAGACAGAGGCGGAAGACATTTGCCAAGCCCCGCTCAC 2228
2223 CATCACCCCGAGCTCCATGAGGAAATCTGCACATCTTCCATAGAAAGTCCCAAGGAGGA 2282
2229 CATCACCCCGAGCTCCATGAGGAAATCTGCACATCTTCCATAGAAAGTCCCAAGGAGGA 2288
2283 CTTCTGTGGTCTCGAAACCACTCAACAGAAATATAGATTTCTAATCTGAGTGTACTGAGC 2342
2289 CTTCTGTGGTCTCGAAACCACTCAACAGAAATATAGATTTCTAATCTGAGTGTACTGAGC 2348
2343 TTTGGTCCACTAAACAAAGCTGAGCTTTGGTCCACTAAACAAAGATGAAATATCAAGAG 2402
2349 TTTGGTCCACTAAACAAAGCTGAGCTTTGGTCCACTAAACAAAGATGAAATATCAAGAG 2408
2403 TGACTCTATAACTCTGTGTCTTTAGAAAGGCTGCCTTTTTCATTTTATAGACAAATCTTTTC 2462
2409 TGACTCTATAACTCTGTGTCTTTAGAAAGGCTGCCTTTTTCATTTTATAGACAAATCTTTTC 2468
2463 AACGCTGAAATGTACTTAATCTGGTCTTACTACCAATATGTATATGATGAGCTTCCCGAGGA 2522
2469 AACGCTGAAATGTACTTAATCTGGTCTTACTACCAATATGTATATGATGAGCTTCCCGAGGA 2528
2523 TGAATGCTGTGTTTAAATTTTCAATAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAG 2582
2529 TGAATGCTGTGTTTAAATTTTCAATAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAG 2588
2583 TCCTTCACTTTTAAATTTTCACTCTCTATAATAATGACATCCCAAGTTCATCGAGGCA 2642
2589 TCCTTCACTTTTAAATTTTCACTCTCTATAATAATGACATCCCAAGTTCATCGAGGCA 2648
2643 AAAAACAAAGTTTCTGTATCTCTGAAACTTTCTATGCTCAGTGGAAGTATCTGCCAGCC 2702
2649 AAAAACAAAGTTTCTGTATCTCTGAAACTTTCTATGCTCAGTGGAAGTATCTGCCAGCC 2708
2703 ACAGATGAGGCCCTGTGAAGCTGACTGAGAAATCTCTGCTGGAAGCCCTGCTCTGT 2762
2709 ACAGATGAGGCCCTGTGAAGCTGACTGAGAAATCTCTGCTGGAAGCCCTGCTCTGT 2768
2763 TCTCCCTCCCAACATGTATAATTTTATTTGAAATACATAATCTTTTCACTATGAAAAAAA 2822
2769 TCTCCCTCCCAACATGTATAATTTTATTTGAAATACATAATCTTTTCACTATGAAAAAAA 2828
2823 AAAAAA 2831
2829 AAAAAA 2837

RESULT 4

AAH72681

ID AAH72681 standard; cDNA; 4422 BP.

XX AC AAH72681;

XX DT 19-SEP-2001 (first entry)

XX DE Human cervical cancer marker nucleic acid 3955.

XX KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

XX Homo sapiens.
OS WO200142467-A2.
PN
XX
XX
PD 14-JUN-2001.
PF
XX
XX 08-DEC-2000; 2000WO-US033312.
XX
XX 08-DEC-1999; 99US-0169681P.
PR 21-DEC-1999; 99US-0171350P.
PR 14-MAR-2000; 2000US-0189315P.
PR 12-MAY-2000; 2000US-0203791P.
PR 09-JUN-2000; 2000US-0210600P.
PR 21-JUL-2000; 2000US-0220114P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
XX
XX Schlegel R, Deeds J, Berger A, Zhao X;
PI
XX WPI; 2001-375006/39.
DR
XX
XX New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.
PT
XX
XX Claim 1; Page 796; 1051pp; English.
PS
XX
XX The invention relates to novel genes (AHU68727-AHU73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy
XX
SQ Sequence 4422 BP; 1285 A; 929 C; 880 G; 1311 T; 0 U; 17 Other;

Query Match 99.0%; Score 2804; DB 4; Length 4422;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2807; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CACGAGCGGAGCTTGGAGCGATTAACGATTGTTGTGTGAGAGGCGCAACGTCGCAATTC 62
DB |||
DB 8 CAGTGGCGGAGCTTGGAGCGATTAACGATTGTTGTGAGAGGCGCAACGTCGCAATTC 67
QY 63 TCCTGAATCTGAGGAGATTTCTACGATTTCTCTCAGCTGAGGCTTTTCCTCCGACCT 122
DB |||
DB 68 TCGTGAATCTGAGGAGATTTCTACGATTTCTCTCAGCTGAGGCTTTTCCTCCGACCT 127
QY 123 GATGCTCTCAATTCGGTCTCGGCGAGCCGAGCTTGGCGTCTGAGAAATGGATGTC 182
DB |||
DB 128 GATGCTCTCAATTCGGTCTCGGCGAGCCGAGCTTGGCGTCTGAGAAATGGATGTC 187
QY 183 TTCAATATACCTCTTCAATTCCTGACTGTTATCAGTCAGTGGTAAATGATGAACA 242
DB |||
DB 188 TTCAATATACCTCTTCAATTCCTGACTGTTATCAGTCAGTGGTAAATGATGAACA 247
QY 243 CACTTCTTATGAGAAACAGGAGTCCAGTTCCTCTCTTTTGGATGATACCTCTCTCTGC 302
DB |||
DB 248 CACTTCTTATGAGAAACAGGAGTCCAGTTCCTCTCTTTTGGATGATACCTCTCTCTGC 307
QY 303 TCCCAATATGGAACATGACTAGCAGTTCGCAATGAGAAAGCTTTGTCGATTTGATAA 362
DB |||
DB 308 TCCCAATATGGAACATGACTAGCAGTTCGCAATGAGAAAGCTTTGTCGATTTGATAA 367
QY 363 CACAGAATCACAAAGTTTCAGAAAGAGTCTTCAAGAAATGATGGCTCACTGGAATGC 422
DB |||
DB 368 CACAGATCACAAAGTTTCAGAAAGAGTCTTCAAGAAATGATGGCTCACTGGAATGC 427
QY 423 CGTCTTTGACCTGGCTGGGTTCTGGTGAACCTTAACTTTTACAGCAGCAGGATGATCA 482
DB |||
DB 428 CGTCTTTGACCTGGCTGGGTTCTGGTGAACCTTAACTTTTACAGCAGCAGGATGATCA 487

QY 483 AACAGCCAAATTTTGGGACGTAAAGAGCTGGTGGAGCTGATTGGAACATGCAAGGTCATCA 542
DB |||
DB 488 AACAGCCAAATTTTGGGACGTAAAGAGCTGGTGGAGCTGATTGGAACATGCAAGGTCATCA 547
QY 543 ATGACAGCCTCAAGTCAGTTCGCTTTTCTAAAGTTTGGAGAAAGCTGTATTCTGTACGGGTGG 602
DB |||
DB 548 ATGACAGCCTCAAGTCAGTTCGCTTTTCTAAAGTTTGGAGAAAGCTGTATTCTGTACGGGTGG 607
QY 603 AAGAGATGGCAACATTAATGCTCTGGGATACAGGTGCAACAAAAGATGGGTTTATAG 662
DB |||
DB 608 AAGAGATGGCAACATTAATGCTCTGGGATACAGGTGCAACAAAAGATGGGTTTATAG 667
QY 663 GCAAGTGAATCAAAATCAGTGGAGCTCAAAATACCTCAGCAAGCAAAACCTTCAAAACC 722
DB |||
DB 668 GCAAGTGAATCAAAATCAGTGGAGCTCAAAATACCTCAGCAAGCAAAACCTTCAAAACC 727
QY 723 CAAGAAGAAACAGAAATTCAAAAGAGCTTGTCTCTTGTGATTTCCAGCAAGGTTTAC 782
DB |||
DB 728 CAAGAAGAAACAGAAATTCAAAAGAGCTTGTCTCTTGTGATTTCCAGCAAGGTTTAC 787
QY 783 TGTGGTCTCTTTCAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGATTAAT 842
DB |||
DB 788 TGTGGTCTCTTTCAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGATTAAT 847
QY 843 CAAGATATGGGATTTACGTAAGAAATTAATCTGCTTATCGCAAGAACCCATAGCATCAA 902
DB |||
DB 848 CAAGATATGGGATTTACGTAAGAAATTAATCTGCTTATCGCAAGAACCCATAGCATCAA 907
QY 903 GTCTTTCTGTACCCAGGTAGCAGCCTCGAAAACTTTGGATATTCAAGTCTGATTTTGA 962
DB |||
DB 908 GTCTTTCTGTACCCAGGTAGCAGCCTCGAAAACTTTGGATATTCAAGTCTGATTTTGA 967
QY 963 TTCACCTGGCTTACTTTATTGTTGTAATTCACAGACGATTAACATCTACATGTTTAAT 1022
DB |||
DB 968 TTCCACTGGCTTACTTTATTGTTGTAATTCACAGACGATTAACATCTACATGTTTAAT 1027
QY 1023 GACTGGGTTGAAGACTTCTCCAGTGGCTTATTTTCAATGGACACCAAGAACTCTACCTTTTA 1082
DB |||
DB 1028 GACTGGGTTGAAGACTTCTCCAGTGGCTTATTTTCAATGGACACCAAGAACTCTACCTTTTA 1087
QY 1083 TGTAAATCCAGCCTTAGTCCAGATGACCACTTTTGTAGTCAAGTGGCTCAAGTGAAGC 1142
DB |||
DB 1088 TGTAAATCCAGCCTTAGTCCAGATGACCACTTTTGTAGTCAAGTGGCTCAAGTGAAGC 1147
QY 1143 TGCCTACATATGGAAGTCTCCACACCTGCAACCTCTCTACTGTCTCTGGGTCAATTC 1202
DB |||
DB 1148 TGCCTACATATGGAAGTCTCCACACCTGCAACCTCTCTACTGTCTCTGGGTCAATTC 1207
QY 1203 TCAGAGGTACAGTCTGTGCTGTGCTGCTCAATCTGACTTTCACAAAGATTGCTACCTGTTTC 1262
DB |||
DB 1208 TCAGAGGTACAGTCTGTGCTGTGCTGCTCAATCTGACTTTCACAAAGATTGCTACCTGTTTC 1267
QY 1263 TGATGACAATACACTTAAATAATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACCAAGAGG 1322
DB |||
DB 1268 TGATGACAATACACTTAAATAATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACCAAGAGG 1327
QY 1323 TGATAAACTTTCCACGGTGGGCTCTCAGAGAAAGAGTCAAGACCTTGGGCT 1382
DB |||
DB 1328 TGATAAACTTTCCACGGTGGGCTCTCAGAGAAAGAGTCAAGACCTTGGGCT 1387
QY 1383 AGTAACAGTAAAGAGTAAAGAGTACTCTGCTCCAAAGCCGCCAGGGTAAAGTGCAATCC 1442
DB |||
DB 1388 AGTAACAGTAAAGAGTAAAGAGTACTCTGCTCCAAAGCCGCCAGGGTAAAGTGCAATCC 1447
QY 1443 ATCCAAATTTCTCCCGCTCATCCGAGCTTGTGCTCCCAAGCTGTGTGGAGACCTTCCCTCT 1502
DB |||
DB 1448 ATCCAAATTTCTCCCGCTCATCCGAGCTTGTGCTCCCAAGCTGTGTGGAGACCTTCCCTCT 1507
QY 1503 TCCTTTCAAAATCTCCTAGTTCCTTAATAAACTCTCTCTGCGCAAGGCCCGGTCTCCCAT 1562
DB |||
DB 1508 TCCTTTCAAAATCTCCTAGTTCCTTAATAAACTCTCTCTGCGCAAGGCCCGGTCTCCCAT 1567

QY 1563 CAACAGAGAGGCTGTGTCTCTCCGCTCTCTCCCAAGCCACCTTCATCTTTCAGATGTC 1622
DB |||||
QY 1568 CAACAGAGAGGCTGTGTCTCTCCGCTCTCTCCCAAGCCACCTTCATCTTTCAGATGTC 1627
DB |||||
QY 1623 GATTAGAACTGGGTGACCCGCAACACCTTCTCATCACCCCATCACTCCACCTGCTTC 1682
DB |||||
QY 1628 GATTAGAACTGGGTGACCCGCAACACCTTCTCATCACCCCATCACTCCACCTGCTTC 1687
DB |||||
QY 1683 GGAGACCAAGATCATGTCTCCGAGAAAAGCCCTTATCTCTGTGAGCCAGAGTCAATCCCA 1742
DB |||||
QY 1688 GGAGACCAAGATCATGTCTCCGAGAAAAGCCCTTATCTCTGTGAGCCAGAGTCAATCCCA 1747
DB |||||
QY 1743 AGCAGAGGCTCTCTGAGTCTAGAAATAGAGTAAGAGGAGGCTAGACTCAAGCTGCT 1802
DB |||||
QY 1748 AGCAGAGGCTCTCTGAGTCTAGAAATAGAGTAAGAGGAGGCTAGACTCAAGCTGCT 1807
DB |||||
QY 1803 GGAGAGTGTGAAAACAAAAGTGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGTATGGCCA 1862
DB |||||
QY 1808 GGAGAGTGTGAAAACAAAAGTGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGTATGGCCA 1867
DB |||||
QY 1863 AGTTGAAAATCTTCAATTTGATCTGTGTGCTTCTGCTGTAACCAAGAGACCTTAGTAA 1922
DB |||||
QY 1868 AGTTGAAAATCTTCAATTTGATCTGTGTGCTTCTGCTGTAACCAAGAGACCTTAGTAA 1927
DB |||||
QY 1923 GGACTCTTAGTCTTACCAATCAAGCAAAATTCGAAGAGCTGTGTACAGTATCTCAGA 1982
DB |||||
QY 1928 GGACTCTTAGTCTTACCAATCAAGCAAAATTCGAAGAGCTGTGTACAGTATCTCAGA 1987
DB |||||
QY 1983 GGCTCGTCTCTATCAGTCCGTGTGTGTGAGAGCTGTGGAAGCTACCTCTTCTCTTT 2042
DB |||||
QY 1988 GGCTCGTCTCTATCAGTCCGTGTGTGAGAGCTGTGGAAGCTACCTCTTCTCTTT 2047
DB |||||
QY 2043 GAGACTTGTGTGAGAGGCTGTGAAATGTGTAGGCAAGAGATAGTTCCTCCAGAGATPAA 2102
DB |||||
QY 2048 GAGACTTGTGTGAGAGGCTGTGAAATGTGTAGGCAAGAGATAGTTCCTCCAGAGATPAA 2107
DB |||||
QY 2103 AAAGTGTGTGTGGCCATGGGAGCCAAACGGAGAGCTGAGATCCATCTCCACGAGTCC 2162
DB |||||
QY 2108 AAAGTGTGTGTGGCCATGGGAGCCAAACGGAGAGCTGAGATCCATCTCCACGAGTCC 2167
DB |||||
QY 2163 GTCATCCAGACACCAATTTCCAGAGAGCAGAGCGGAAGACATTCGAAGCCCGGTCCAC 2222
DB |||||
QY 2168 GTCATCCAGACACCAATTTCCAGAGAGCAGAGCGGAAGACATTCGAAGCCCGGTCCAC 2227
DB |||||
QY 2223 CATCAGCCAGCTCCATGAGAGAAATCTGCACATATTCATAGAAAGTCCAGGAGGA 2282
DB |||||
QY 2228 CATCAGCCAGCTCCATGAGAGAAATCTGCACATATTCATAGAAAGTCCAGGAGGA 2287
DB |||||
QY 2283 CTTCTGTGTCTGACACTCAACAGAAATATAGATTTCTATCTGAGTGAGTTACTGAGC 2342
DB |||||
QY 2288 CTTCTGTGTCTGACACTCAACAGAAATATAGATTTCTATCTGAGTGAGTTACTGAGC 2347
DB |||||
QY 2343 TTTGTCTCACTAAAACAGCTGAGCTTTGGTCCACTAAAACAGATGAAAATACAGAG 2402
DB |||||
QY 2348 TTTGTCTCACTAAAACAGCTGAGCTTTGGTCCACTAAAACAGATGAAAATACAGAG 2407
DB |||||
QY 2403 TGACTCTAATCTGGTCTTTAAGAAAGCTGCCCTTTTCAATTTTATAGAAAATCTTTTC 2462
DB |||||
QY 2408 TGACTCTAATCTGGTCTTTAAGAAAGCTGCCCTTTTCAATTTTATAGAAAATCTTTTC 2467
DB |||||
QY 2463 AACGCTGAATGTACTTAATCTGGTCTTACTACCAATATGATATGACGCTTCCCGAGGA 2522
DB |||||
QY 2468 AACGCTGAATGTACTTAATCTGGTCTTACTACCAATATGATATGACGCTTCCCGAGGA 2527
DB |||||
QY 2523 TGAATGCTGTGTTTAAATTTTCAAAAGTAAATTTGTCTACTCTAGCATTTTGAATGAATAG 2582
DB |||||
QY 2583 TCTTCACTTTTAAATTTATCTTCTCTATAATAATGACATCCAGTTCATGAGGCA 2642
DB |||||
QY 2588 TCTTCACTTTTAAATTTATCTTCTCTATAATAATGACATCCAGTTCATGAGGCA 2647
DB |||||
QY 2643 ABAACACAGTTCTTGTGTTATCTGAAACTTTTCTATGCTCAGTGGAAAGTATCTGCCAGCC 2702
DB |||||

DB 2648 AAAAACAAGTTTCTTGTATCTCTGAAACTTTTATGTCTAGTGAAGATATCTGCCAGCC 2707
QY 2703 ACAGCATGAGGCTGTGAAGGCTGACTGAGAAATCTCTCTGCTGAAGACCCCTGTTCTGT 2762
DB 2708 ACAGCATGAGGCTGTGAAGGCTGACTGAGAAATCTCTCTGCTGAAGACCCCTGTTCTGT 2767
QY 2763 TCTGCTCCCAACATGATTAATTTTATTTGAAATACATAATCTTTTCACTATG 2814
DB 2768 TCTGCTCCCAACATGATTAATTTTATTTGAAATACATAATCTTTTCACTATG 2819
RESULT 5
ABV25710
ID ABV25710 standard; cDNA; 4422 BP.
XX
AC ABV25710;
XX
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 25701.
XX
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200160860-A2.
FN
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US0005171.
XX
XX 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
PI
XX
DR WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.
PT
PT
XX
PS Claim 1; Page 5158-5159; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the efficacy progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 4422 BP; 1285 A; 929 C; 880 G; 1311 T; 0 U; 17 Other;
Query Match 99.0%; Score 2804; DB 5; Length 4422;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2807; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 3 CACGAGCGGAGTTGGAGGCGATAAACGATTTGTTGTGAGAGCGCAACCTGCGATTTC 62
||| |||||

Db 8 CAGTGGCGGAGTTGGAGGCGATAACGATTTGTGTTGTGAGAGGCGCAAGCTGCGATTTC 67
Qy 63 TGCTGAACCTTGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCT 122
Db 68 TGCTGAACCTTGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCT 127
Qy 123 GATGCTCTTCAATTCGGTGCTCGCGCAGCCGAGCTTTGGCGTCTTGAGAAATGGATGGTC 182
Db 128 GATGCTCTTCAATTCGGTGCTCGCGCAGCCGAGCTTTGGCGTCTTGAGAAATGGATGGTC 187
Qy 183 TTCAAAATACCTCTTCAATCCCTTCTGACTGGTTATCAGTCAGTGGTGAATGATGAACA 242
Db 188 TTCAAAATACCTCTTCAATCCCTTCTGACTGGTTATCAGTCAGTGGTGAATGATGAACA 247
Qy 243 CACTTCTTATGAGAAACAGGAGTCCAGTCTCTCTTTTGGATGATCTCTCTCTTCGC 302
Db 248 CACTTCTTATGAGAAACAGGAGTCCAGTCTCTCTCTTTTGGATGATCTCTCTCTTCGC 307
Qy 303 TCCCAATATGGAACATGTACTAGCAGTTGCCAATGAAGAAGGCTTTGTTCCGATTGTATAA 362
Db 308 TCCCAATATGGAACATGTACTAGCAGTTGCCAATGAAGAAGGCTTTGTTCCGATTGTATAA 367
Qy 363 CACAGAATCACAAAGTTTCAGAAAGAGTGTCTTCAAAAGAAATGGATGGTCTCACTGGAATGC 422
Db 368 CACAGAATCACAAAGTTTCAGAAAGAGTGTCTTCAAAAGAAATGGATGGTCTCACTGGAATGC 427
Qy 423 CGTCTTTGACCTGGCTGGGTTCTGGTGAACTTAAACTTTGTTACAGCAGCAGGTGATCA 482
Db 428 CGTCTTTGACCTGGCTGGGTTCTGGTGAACTTAAACTTTGTTACAGCAGCAGGTGATCA 487
Qy 483 AACAGCCAAATTTTGGAGCTTAAAGCTGGTGAGCTGATTGGAAATGCAAAAGGTGATCA 542
Db 488 AACAGCCAAATTTTGGAGCTTAAAGCTGGTGAGCTGATTGGAAATGCAAAAGGTGATCA 547
Qy 543 ATGAGCTTCAAGTCAAGTTCCTTTTCTAAAGTTTGAGAAAGCTGATTTCTGACGGGTGG 602
Db 548 ATGAGCTTCAAGTCAAGTTCCTTTTCTAAAGTTTGAGAAAGCTGATTTCTGACGGGTGG 607
Qy 603 AAGAGTGGCAATATATGCTGGATACAGGTGCAACAAAGATGGGTTTTATAG 662
Db 608 AAGAGTGGCAATATATGCTGGATACAGGTGCAACAAAGATGGGTTTTATAG 667
Qy 663 GCAAGTGAATCAAACTCAGTGGAGCTCACAAATACCTCAGACAAGCAACCCCTTCAAACC 722
Db 668 GCAAGTGAATCAAACTCAGTGGAGCTCACAAATACCTCAGACAAGCAACCCCTTCAAACC 727
Qy 723 CAAGAAGAAACAGAAATTCAAAAGGAGCTGCTCCTCTCTGTGGATTTCCAGCAAGGTGTAC 782
Db 728 CAAGAAGAAACAGAAATTCAAAAGGAGCTGCTCCTCTCTGTGGATTTCCAGCAAGGTGTAC 787
Qy 783 TGTGGTCTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGCAGCTGTGGATGGGATAT 842
Db 788 TGTGGTCTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGCAGCTGTGGATGGGATAT 847
Qy 843 CAAAGTATGGGATTTACGTAAGAAATATATCTGCTTATCGACAAGAACCCATAGCATCCAA 902
Db 848 CAAAGTATGGGATTTACGTAAGAAATATATCTGCTTATCGACAAGAACCCATAGCATCCAA 907
Qy 903 GTCTTTCTGTATCCAGGTAGCAGCATCTCGAAAACCTTGGATATTCAGGCTGATTTTGA 962
Db 908 GTCTTTCTGTATCCAGGTAGCAGCATCTCGAAAACCTTGGATATTCAGGCTGATTTTGA 967
Qy 963 TTCCACTGGCTCTACTTTATTTGCTAAATTTGGACAGCATATCATCTACATGTTTAATAT 1022
Db 968 TTCCACTGGCTCTACTTTATTTGCTAAATTTGGACAGCATATCATCTACATGTTTAATAT 1027
Qy 1023 GACTGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACCGAACTCTACCTTTTA 1082
Db 1028 GACTGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACCGAACTCTACCTTTTA 1087
Qy 1083 TGTAAATCCAGCTTAGTCTCAGATGACCAAGTCTTGTAGTCAAGTGGCTCAAGTGAAGC 1142
Db 1088 TGTAAATCCAGCTTAGTCTCAGATGACCAAGTCTTGTAGTCAAGTGGCTCAAGTGAAGC 1147

Qy 1143 TGCCTACATATGGAAGGTCTCCACACCTGCGCAACCTCTACTGTGTCTCTGGGTCAATC 1202
Db 1148 TGCCTACATATGGAAGGTCTCCACACCTGCGCAACCTCTACTGTGTCTCTGGGTCAATC 1207
Qy 1203 TCAAGAGGTCAAGTCTGTGTGCTGGTGTCCATCTGACTTCAAAAGATTTGCTGATCTTC 1262
Db 1208 TCAAGAGGTCAAGTCTGTGTGCTGGTGTCCATCTGACTTCAAAAGATTTGCTGATCTTC 1267
Qy 1263 TGATGACAAATACATCAAAATCTGGCGCTTGAATAGAGGCTTAGAGGAAACAGGAGG 1322
Db 1268 TGATGACAAATACATCAAAATCTGGCGCTTGAATAGAGGCTTAGAGGAAACAGGAGG 1327
Qy 1323 TGATAAACTTTTCCAGCGTGGGTGGGCTCTCAGAGAAAAAGAGTCAAGACCTTGGCT 1382
Db 1328 TGATAAACTTTTCCAGCGTGGGTGGGCTCTCAGAGAAAAAGAGTCAAGACCTTGGCT 1387
Qy 1383 AGTAACAGTAAACGAGTAGCAGAGTACTCTGCGCAAGCCCTCAGAGGCTTGAAGTCAATTC 1442
Db 1388 AGTAACAGTAAACGAGTAGCAGAGTACTCTGCGCAAGCCCTCAGAGGCTTGAAGTCAATTC 1447
Qy 1443 ATCCAAATTTTCCCGCTCATCCGAGCTTGTGCGCAAGCTGTCTGGAGACCTCCCTCT 1502
Db 1448 ATCCAAATTTTCCCGCTCATCCGAGCTTGTGCGCAAGCTGTCTGGAGACCTCCCTCT 1507
Qy 1503 TCCTTCAAAATCTCTACGTTCTCTATTAAAACTCTCTGCGCAAGCCCTCTCCCAT 1562
Db 1508 TCCTTCAAAATCTCTACGTTCTCTATTAAAACTCTCTGCGCAAGCCCTCTCCCAT 1567
Qy 1563 CAACAGAGAGGCTCTGTCTCTCTCTCTCCCAAGCCCTCTCTTTCAAGATGTC 1622
Db 1568 CAACAGAGAGGCTCTGTCTCTCTCTCTCTCCCAAGCCCTCTCTTTCAAGATGTC 1627
Qy 1623 GATTAGAACTGGGTGAGCCCGAACACCTTCTCTCATCACCACCTCATCCACCTGCTTC 1682
Db 1628 GATTAGAACTGGGTGAGCCCGAACACCTTCTCTCATCACCACCTCATCCACCTGCTTC 1687
Qy 1683 GGAGACCAAGATCATGTCTCGAGAAAGCCCTTATTCTGTGAGCCAGAGTCAATCCA 1742
Db 1688 GGAGACCAAGATCATGTCTCGAGAAAGCCCTTATTCTGTGAGCCAGAGTCAATCCA 1747
Qy 1743 AGCAGAGGCTTGTCTGAGTCTAGAAATAGAGTAAAGAGGAGGTAGACTCAAGCTGTCT 1802
Db 1748 AGCAGAGGCTTGTCTGAGTCTAGAAATAGAGTAAAGAGGAGGTAGACTCAAGCTGTCT 1807
Qy 1803 GGAGAGTGTGAAAACAAAGTGTGAAAGTGTAACTGTGTGACTGAGCTTGTAGGCCA 1862
Db 1808 GGAGAGTGTGAAAACAAAGTGTGAAAGTGTAACTGTGTGACTGAGCTTGTAGGCCA 1867
Qy 1863 AGTTGMAAATCTTCAATTTGGATCTGTGCTGCTTGTGTTAAACAGGAAGACCTTAGTAA 1922
Db 1868 AGTTGMAAATCTTCAATTTGGATCTGTGCTGCTTGTGTTAAACAGGAAGACCTTAGTAA 1927
Qy 1923 GGACTCTCTAGGCTCTACCAAAATCAAGCAAAATTTGAAGAGCTGGTACAGTATCTCAGA 1982
Db 1928 GGACTCTCTAGGCTCTACCAAAATCAAGCAAAATTTGAAGAGCTGGTACAGTATCTCAGA 1987
Qy 1983 GCCTCGCTCTCTACGTCCGTATGCTTCAAGAAAGCTGTGGAAGCTTACCTCTCTCTTT 2042
Db 1988 GCCTCGCTCTCTACGTCCGTATGCTTCAAGAAAGCTGTGGAAGCTTACCTCTCTCTTT 2047
Qy 2043 GAGACCTTGTGAGAGGCTCTGAAATGGTAGGCAAGAGAAATAGTTCCTCCAGAGAAATAA 2102
Db 2048 GAGACCTTGTGAGAGGCTCTGAAATGGTAGGCAAGAGAAATAGTTCCTCCAGAGAAATAA 2107
Qy 2103 AAACTGTTGTTGGCCATGGCAGCCAAACGGAAGGCTGAGAAATCCATCTCCACGAAGTCC 2162
Db 2108 AAACTGTTGTTGGCCATGGCAGCCAAACGGAAGGCTGAGAAATCCATCTCCACGAAGTCC 2167
Qy 2163 GTCTATCCAGACACCCAAATTCAGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTAC 2222
Db 2168 GTCTATCCAGACACCCAAATTCAGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTAC 2227

QY 2223 CATCAGCCAGCTCCATGAGGAAATCTGCACATACCTTCCATAGAAAGTCCAGGAGGA 2282
Db |||||
QY 2228 CATCAGCCAGCTCCATGAGGAAATCTGCACATACCTTCCATAGAAAGTCCAGGAGGA 2287
Db |||||
QY 2283 CTTCTGTGCTCTGAACACTCAACAGAAATTATAGATTCTTAATCTCAGTGAGTTACTGAGC 2342
Db |||||
QY 2288 CTTCTGTGCTCTGAACACTCAACAGAAATTATAGATTCTTAATCTCAGTGAGTTACTGAGC 2347
Db |||||
QY 2343 TTTGGTCCATAAAACAAGCTGAGCTTTGGTCCATAAAACAAGATGAAAAATACAAGAG 2402
Db |||||
QY 2348 TTTGGTCCATAAAACAAGCTGAGCTTTGGTCCATAAAACAAGATGAAAAATACAAGAG 2407
Db |||||
QY 2403 TGACTCTATAACTCTGGTCTTTAAGAAAGCTGCGCTTTTCAATTTTGTAGCAAAATCTTTTC 2462
Db |||||
QY 2408 TGACTCTATAACTCTGGTCTTTAAGAAAGCTGCGCTTTTCAATTTTGTAGCAAAATCTTTTC 2467
Db |||||
QY 2463 AACGCTGAAATGTACCTAATCTGGTTCTACTACCAATATGATATGACGCTTCCCGAGGA 2522
Db |||||
QY 2468 AACGCTGAAATGTACCTAATCTGGTTCTACTACCAATATGATATGACGCTTCCCGAGGA 2527
Db |||||
QY 2523 TGAATGCTGTGTTAAATTTCAATAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAG 2582
Db |||||
QY 2528 TGAATGCTGTGTTAAATTTCAATAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAG 2587
Db |||||
QY 2583 TCTTCACTTTTAAATTTATCATCTCTCTAATAATGACATCCAGTTCAATGAGGGA 2642
Db |||||
QY 2588 TCTTCACTTTTAAATTTATCATCTCTCTAATAATGACATCCAGTTCAATGAGGGA 2647
Db |||||
QY 2643 AAAAAACAAGTTTCTGTTATCTCTGAAACTTTTCTATGCTCAGTGGAAGTATCTGCCAGCC 2702
Db |||||
QY 2648 AAAAAACAAGTTTCTGTTATCTCTGAAACTTTTCTATGCTCAGTGGAAGTATCTGCCAGCC 2707
Db |||||
QY 2703 ACAGCATGAGGCTGTGAAGGCTGACTGAGAAATCCTCTGCTGAAAGACCCCTGGTTCTGT 2762
Db |||||
QY 2708 ACAGCATGAGGCTGTGAAGGCTGACTGAGAAATCCTCTGCTGAAAGACCCCTGGTTCTGT 2767
Db |||||
QY 2763 TCTGCTCCAACTGATATATTTTATTTTGAATACATATCTTTTCACTATG 2814
Db |||||
QY 2768 TCTGCTCCAACTGATATATTTTATTTTGAATACATATCTTTTCACTATG 2819
Db |||||

RESULT 6

ADL61974
ID ADL61974 standard; DNA; 4422 BP.

XX AC ADL61974;

XX DT 20-MAY-2004 (first entry)

XX DE Human ovarian cancer DNA marker #20186.

XX KW Human; ovarian cancer; ds; tumour; cytostatic; DNA marker.

XX OS Homo sapiens.

XX PN WO200170979-A2.

XX PD 27-SEP-2001.

XX PF 21-MAR-2001; 2001WO-US009126.

XX PR 21-MAR-2000; 2000US-0191031P.

XX PR 25-MAY-2000; 2000US-0207124P.

XX PR 15-JUN-2000; 2000US-0211940P.

XX PR 07-JUL-2000; 2000US-02116820P.

XX PR 25-JUL-2000; 2000US-0220661P.

XX PR 21-DEC-2000; 2000US-0257672P.

XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Lee J, Lillie J;

XX WPI; 2001-611502/70.

XX Novel isolated nucleic acid molecules (markers) overexpressed in ovarian
PT cancer cells as compared to their normal non-cancerous ovarian cells are
PT used to characterize stage, grade, histological type of ovarian cancer.
XX Disclosure; SEQ ID NO 20186; 106pp; English.

XX The invention relates to nucleic acid markers which are overexpressed in
CC ovarian cancer cells as compared to their expression in normal (i.e. non-
CC cancerous) ovarian cells. The invention also relates to polypeptides
CC encoded by the markers, antibodies that selectively bind to the
CC polypeptides, a method of inhibiting ovarian cancer in a patient at risk
CC of developing ovarian cancer involving inhibiting expression of a gene
CC corresponding to a marker of the invention and a method of treating a
CC patient afflicted with ovarian cancer comprising providing to cells of
CC the patient an antisense oligonucleotide complementary to a marker of the
CC invention. The markers are useful for assessing if a patient is afflicted
CC with ovarian cancer, which involves comparing the level of expression of
CC a marker in a patient sample and a normal level of expression of the
CC marker in a control non-ovarian cancer sample. A difference between the
CC expression levels indicates ovarian cancer. The level of expression of a
CC marker corresponds to a secreted protein or to a transcribed
CC polynucleotide or its portion. The level of expression of the marker is
CC assessed by detecting the presence in the sample, a protein or protein
CC fragment corresponding to the marker. The presence of protein or protein
CC fragment is detected using an antibody that specifically binds with the
CC protein or protein fragment. Alternatively, the level of expression of
CC the marker is assessed by detecting the presence of a transcribed
CC polynucleotide which anneals with the marker or anneals with a portion of
CC the polynucleotide comprising the marker, under stringent conditions. The
CC marker is also used for monitoring the progression of ovarian cancer in a
CC patient which involves detecting expression of the marker in a patient
CC sample at a first point in time, repeating the method at a subsequent
CC time and comparing the level of expression. The method is carried out
CC using an ovarian tissue sample. A composition comprising a marker,
CC polypeptide or antibody of the invention is used to treat ovarian cancer.
CC This sequence represents a human ovarian cancer DNA marker of the
CC invention.

XX SQ Sequence 4422 BP; 1285 A; 929 C; 880 G; 1311 T; 0 U; 17 Other;

Query Match 99.0%; Score 2804; DB 5; Length 4422;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2807; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CACGAGCGGAGTTGGAGGCGATTAACGATTTGTTGTGAGAGCGCAACGTCGATTTC 62
Db |||||
8 CAGTGGCGGAGTTGGAGGCGATTAACGATTTGTTGTGAGAGCGCAACGTCGATTTC 67

QY 63 TGCTGAACCTGGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTCTCTCCGACCCCT 122
Db |||||
68 TGCTGAACCTGGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTCTCTCCGACCCCT 127

QY 123 GATGCTCTTCAATTCGGTGTCTCCGCCAGCCAGCTTGGCGTCTCTGAGAAATGGATGTC 182
Db |||||
128 GATGCTCTTCAATTCGGTGTCTCCGCCAGCCAGCTTGGCGTCTCTGAGAAATGGATGTC 187

QY 183 TTCACAATACCCCTCTTCAATCCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACA 242
Db |||||
188 TTCACAATACCCCTCTTCAATCCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACA 247

QY 243 CACTTCTTATGGAGAAACAGAGTCCAGTTCCTCTCTTGGATGTACCTTCTCTCTGCG 302
Db |||||
248 CACTTCTTATGGAGAAACAGAGTCCAGTTCCTCTCTTGGATGTACCTTCTCTCTGCG 307

QY 303 TCCCAATATGGAACATGTACTAGCAGTTTGCCAAATGAAGAGGCTTTTGTTCGATTGTATAA 362
Db |||||
308 TCCCAATATGGAACATGTACTAGCAGTTTGCCAAATGAAGAGGCTTTTGTTCGATTGTATAA 367

QY 363 CACAGAATCACAAGTTTTCAGAAAGAGAGTGGTTCCAAAGATGGATGGCTCACTGGAATGC 422
Db |||||
368 CACAGAATCACAAGTTTTCAGAAAGAGAGTGGTTCCAAAGATGGATGGCTCACTGGAATGC 427

QY 423 CGTCTTTGACGTGGCTGGTTCCTGGTGAACTTAAACTTTGTACAGCAGCAGGTGATCA 482
Db 428 CGTCTTTGACCTGGCTGGTTCCTGGTGAACTTAAACTTTGTACAGCAGCAGGTGATCA 487
QY 483 AACAGCAAATTTTGGGACGTAAAGCTGGTGAGCTGATGTGAACATGCAAAAGGTGATCA 542
Db 488 AACAGCAAATTTTGGGACGTAAAGCTGGTGAGCTGATGTGAACATGCAAAAGGTGATCA 547
QY 543 ATGCAGCTCAAGTCAAGTTCCTTTCTAAAGTTTGAGAAAGCTGTATTTCTGTACGGGTGG 602
Db 548 ATGCAGCTCAAGTCAAGTTCCTTTCTAAAGTTTGAGAAAGCTGTATTTCTGTACGGGTGG 607
QY 603 AAGAGATGCAACATTTATGCTGCGGATACAGGTGCAACAAAAGATGGGTTTATAG 662
Db 608 AAGAGATGCAACATTTATGCTGCGGATACAGGTGCAACAAAAGATGGGTTTATAG 667
QY 663 GCAAGTGAATCAAAATCAGTGGAGCTCAAAATACCTCAGACAAGCAAAACCCCTTCAAAACC 722
Db 668 GCAAGTGAATCAAAATCAGTGGAGCTCAAAATACCTCAGACAAGCAAAACCCCTTCAAAACC 727
QY 723 CAAGAAGAAACAGAAATTCAAAAGGACTTGCTCCTCTCTGTGGATTTCCAGCAAGGTGTAC 782
Db 728 CAAGAAGAAACAGAAATTCAAAAGGACTTGCTCCTCTCTGTGGATTTCCAGCAAGGTGTAC 787
QY 783 TGTGGTCTCTTTCAAGACAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAT 842
Db 788 TGTGGTCTCTTTCAAGACAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAT 847
QY 843 CAAAGTATGGGATTTACGTAAGAAATATATCTGCTTATCGACAAGAAACCCATAGCATCAA 902
Db 848 CAAAGTATGGGATTTACGTAAGAAATATATCTGCTTATCGACAAGAAACCCATAGCATCAA 907
QY 903 GTCTTTCTGTACCGAGTACAGACATCGAAGAACTTGGATATTCAGTCTGATTTTGA 962
Db 908 GTCTTTCTGTACCGAGTACAGACATCGAAGAACTTGGATATTCAGTCTGATTTTGA 967
QY 963 TTCCACTGGCTCTACTTTATTTGCTAAATGACAGACGATACATCTACATGTTTAAAT 1022
Db 968 TTCCACTGGCTCTACTTTATTTGCTAAATGACAGACGATACATCTACATGTTTAAAT 1027
QY 1023 GACTGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACCGAACTCTACCTTTTA 1082
Db 1028 GACTGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACCGAACTCTACCTTTTA 1087
QY 1083 TGTAAATCCAGCTTAGTCCAGATGACAGATTTTGTAGTGGCTCAAGTGATGAAGC 1142
Db 1088 TGTAAATCCAGCTTAGTCCAGATGACAGATTTTGTAGTGGCTCAAGTGATGAAGC 1147
QY 1143 TGCCTACATATGGAAGGTCTCCACACCCCTGGCAACCTCTACTGTGCTCCTGGGTCAATC 1202
Db 1148 TGCCTACATATGGAAGGTCTCCACACCCCTGGCAACCTCTACTGTGCTCCTGGGTCAATC 1207
QY 1203 TCAAGAGGTCAAGTCTGTGTGTGCTGCTCCATCTGACTTCAAAAGATTTGCTACCTGTC 1262
Db 1208 TCAAGAGGTCAAGTCTGTGTGTGCTGCTCCATCTGACTTCAAAAGATTTGCTACCTGTC 1267
QY 1263 TGAATGACAAATACATTAATAATCTGGGCTTGAATAGAGCTTAGAGGAAACACAGGAGG 1322
Db 1268 TGAATGACAAATACATTAATAATCTGGGCTTGAATAGAGCTTAGAGGAAACACAGGAGG 1327
QY 1323 TGAATAACTTTTCCAGGTGGTGGGCTCTCAGAGAAAGAGTCAAGACCTGGCT 1382
Db 1328 TGAATAACTTTTCCAGGTGGTGGGCTCTCAGAGAAAGAGTCAAGACCTGGCT 1387
QY 1383 AGTAAACAGTAAACAGTAGCAGATACCTCTGCCAAAGCCCCCAGGGTAAAGTGAATCC 1442
Db 1388 AGTAAACAGTAAACAGTAGCAGATACCTCTGCCAAAGCCCCCAGGGTAAAGTGAATCC 1447
QY 1443 ATCCAAATTTCCCGGTGATCCGAGCTGTGCCCCCAAGCTGTGCTGGAGACCTCCCTCT 1502
Db 1448 ATCCAAATTTCCCGGTGATCCGAGCTGTGCCCCCAAGCTGTGCTGGAGACCTCCCTCT 1507
QY 1503 TCCTTCAAAATACCTCTACGTTCTCTATTAAACCTCTCTCTGCGCAAGGCCGCTCTCCCAT 1562

Db 1508 TCCTTCAAAATACCTCTACGTTCTCTATTAAACCTCTCTCTGCGCAAGGCCGCTCTCCCAT 1567
QY 1563 CAACAGAGAGCTCTGTCTCTCTCGGTCTCTCCCAAGCCACCTTCATCTTTCAAGATGTC 1622
Db 1568 CAACAGAGAGCTCTGTCTCTCTCGGTCTCTCCCAAGCCACCTTCATCTTTCAAGATGTC 1627
QY 1623 GATTAGAAACTGGGTGACCCCGAAACACTTCTCTCATCACACCCCATCACTCCACCTGCTTC 1682
Db 1628 GATTAGAAACTGGGTGACCCCGAAACACTTCTCTCATCACACCCCATCACTCCACCTGCTTC 1687
QY 1683 GGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCAATCCA 1742
Db 1688 GGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCAATCCA 1747
QY 1743 AGCAGAGCTTGGCTCTGAGTCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGCT 1802
Db 1748 AGCAGAGCTTGGCTCTGAGTCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGCT 1807
QY 1803 GGAGAGTGTGAAAACAAAAGTGTGAAAGTGTGTAACCTGTGTGACTGAGCTTGTATGGCCA 1862
Db 1808 GGAGAGTGTGAAAACAAAAGTGTGAAAGTGTGTAACCTGTGTGACTGAGCTTGTATGGCCA 1867
QY 1863 AGTTGAAAATCTTCAATTTGGATCTGTCTGCTCTTGTCTGTTAAACAGGAAGACCTTAGTAA 1922
Db 1868 AGTTGAAAATCTTCAATTTGGATCTGTCTGCTCTTGTCTGTTAAACAGGAAGACCTTAGTAA 1927
QY 1923 GGACTCTAGGCTCTTACCAAAATCAAGCAAAATTTGAAGGAGCTGGTACCAGTATCTAGA 1982
Db 1928 GGACTCTAGGCTCTTACCAAAATCAAGCAAAATTTGAAGGAGCTGGTACCAGTATCTAGA 1987
QY 1983 GCCTCCGCTCTCCTATCAGTCCGTATGTTTCAAAAAGCTGTGGAAGCTTACCTCTTCTCTTT 2042
Db 1988 GCCTCCGCTCTCCTATCAGTCCGTATGTTTCAAAAAGCTGTGGAAGCTTACCTCTTCTCTTT 2047
QY 2043 GAGACCTTTGTGGAGAAGGCTCTGAAATGGTGAAGCAAGAGAAATAGTTCCCCAGAGAAATAA 2102
Db 2048 GAGACCTTTGTGGAGAAGGCTCTGAAATGGTGAAGCAAGAGAAATAGTTCCCCAGAGAAATAA 2107
QY 2103 AAACCTGGTTGTGGCCATGGCAGCCAAACGGAAGGCTGAGAAATCCATCTCCACGAGTCC 2162
Db 2108 AAACCTGGTTGTGGCCATGGCAGCCAAACGGAAGGCTGAGAAATCCATCTCCACGAGTCC 2167
QY 2163 GTCAATCCAGACACCCCAATTTCCAGAGACAGAGCGGAAGACATTTGCCAAGCCGGTCAAC 2222
Db 2168 GTCAATCCAGACACCCCAATTTCCAGAGACAGAGCGGAAGACATTTGCCAAGCCGGTCAAC 2227
QY 2223 CATCACGCCAGCTCCATGAGGAAAATCTGCACATACCTTCATAGAAAAGTCCCAAGGAGGA 2282
Db 2228 CATCACGCCAGCTCCATGAGGAAAATCTGCACATACCTTCATAGAAAAGTCCCAAGGAGGA 2287
QY 2283 CTTCTGTGGTCTGAAACACTCAACAGAAATTAAGATTTCTAATCTGAGTGAAGTACTGAGC 2342
Db 2288 CTTCTGTGGTCTGAAACACTCAACAGAAATTAAGATTTCTAATCTGAGTGAAGTACTGAGC 2347
QY 2343 TTTGCTCCACTAAACAAAGCTGAGCTTGGTCCACTAAAACAAAGATGAAAATACAAGAG 2402
Db 2348 TTTGCTCCACTAAACAAAGCTGAGCTTGGTCCACTAAAACAAAGATGAAAATACAAGAG 2407
QY 2403 TGACTCTATAACTCTGGTCTTTTAAAGAAAGCTGCCCTTTTTCATTTTATAGCAAAATCTTTTC 2462
Db 2408 TGACTCTATAACTCTGGTCTTTTAAAGAAAGCTGCCCTTTTTCATTTTATAGCAAAATCTTTTC 2467
QY 2463 AACGCTGAAATGTACCTAAATCTGGTCTTAACCATTAATGTATGTATGTAGCTTCCGAGGA 2522
Db 2468 AACGCTGAAATGTACCTAAATCTGGTCTTAACCATTAATGTATGTATGTAGCTTCCGAGGA 2527
QY 2523 TGAATGCTGTCTTAAATTTTCAATGAATTAATTTGTCTACTAGCATTTTCAATTTGAATGAATAG 2582
Db 2528 TGAATGCTGTCTTAAATTTTCAATGAATTAATTTGTCTACTAGCATTTTGAATGAATAG 2587
QY 2583 TCCTTCAAAATACCTCTACGTTCTCTCTATAATAATGACATCCAGTTCATGAGGACA 2642

1176	Db	 TGCCCTACATATGAAGGCTCTCCACACCGCTGGCAACCTCTCTACTGTGCTCTGGGTCAATTC	1235
1203	Qy	 TCAAGAGGTCACGTCGTCTGTGTGGTGTGCATCTTGACTTCACAAAGATTGCTACTCTGTTC	1262
1236	Db	 TCAAGAGGTCACGTCGTCTGTGTGGTGTGCATCTTGACTTCACAAAGATTGCTACTGTTC	1295
1263	Qy	 TGATGACAAATACACTAAAAAATCTGGCGCTTTGAATAGAGGCTTAGAGAGAGAAACGAGGAGG	1322
1296	Db	 TGATGACAAATACACTAAAAAATCTGGCGCTTTGAATAGAGGCTTAGAGAGAGAAACGAGGAGG	1355
1323	Qy	 TGATAAACTTTCCACGCTGGGTTGGGCGCTCTCAGAGAGAAAAAGATCAAGACTGGCCCT	1382
1356	Db	 TGATAAACTTTTCCACGCTGGGTTGGGCGCTCTCAGAGAGAAAAAGATCAAGACTGGCCCT	1415
1383	Qy	 AGTAACAGTAACAGTAGTCAGAGTAGTACTCTGTGCCAAAGCCGCCAGGGTAAGTGCATATCC	1442
1416	Db	 AGTAACAGTAACAGTAGTCAGAGTAGTACTCTGTGCCAAAGCCGCCAGGGTAAGTGCATATCC	1475
1443	Qy	 ATCCAAATTTCTTCCC CGTCAATCCGAGCTTTGTGCCCAAGCTGTGCTGGAGACCTCCCTCT	1502
1476	Db	 ATCCAAATTTCTTCCC CGTCAATCCGAGCTTTGTGCCCAAGCTGTGCTGGAGACCTCCCTCT	1535
1503	Qy	 TCCTTCAAAATCTCTCTAGTTTCTCTATTAAAACTCTCTCTGCCAAGGCCCGGTCTCCCAT	1562
1536	Db	 TCCTTCAAAATCTCTCTAGTTTCTCTATTAAAACTCTCTCTGCCAAGGCCCGGTCTCCCAT	1595
1563	Qy	 CAACAGAAAGGCTCTGTCCTCCGTCCTCTCCCAAGCCACTTCATCTTTTCAAGATGTC	1622
1596	Db	 CAACAGAAAGGCTCTGTCCTCCGTCCTCTCCCAAGCCACTTCATCTTTTCAAGATGTC	1655
1623	Qy	 GATTAGAAACTGGGTGACCCGGAACACCTTCTCATACACACCCATCACTCCACTCTGCTTC	1682
1656	Db	 GATTAGAAACTGGGTGACCCGGAACACCTTCTCATACACACCCATCACTCCACTCTGCTTC	1715
1683	Qy	 GGAGACCAAGATCATGTCTCCGAGAAAAGCCCTTATTCTGTGAGCCAGAAAGTCATCCCA	1742
1716	Db	 GGAGACCAAGATCATGTCTCCGAGAAAAGCCCTTATTCTGTGAGCCAGAAAGTCATCCCA	1775
1743	Qy	 AGCAGAGGCTTGCTCTCAGTCTAGAAATAGAGTAAGAGAGAGGCTAGACTCAAGCTGTCT	1802
1776	Db	 AGCAGAGGCTTGCTCTCAGTCTAGAAATAGAGTAAGAGAGAGGCTAGACTCAAGCTGTCT	1835
1803	Qy	 GGAGAGTGTCAAAACAAAAGTGTGAAAGAGTGTGTAACTGTGTGACTGAGCTTGATGGCCA	1862
1836	Db	 GGAGAGTGTCAAAACAAAAGTGTGAAAGAGTGTGTAACTGTGTGACTGAGCTTGATGGCCA	1895
1863	Qy	 AGTTGAAAAATCTTCATTTGGATCTGTCTGCCCTTGCTGGTAAACAGGAAGACCTTAGTAA	1922
1896	Db	 AGTTGAAAAATCTTCATTTGGATCTGTCTGCCCTTGCTGGTAAACAGGAAGACCTTAGTAA	1955
1923	Qy	 GGACTCTCTAGGCTCTACCAAAATCAAGCAAAAATGAAGAGAGCTGGTACCAAGTATCTCAGA	1982
1956	Db	 GGACTCTCTAGGCTCTACCAAAATCAAGCAAAAATGAAGAGAGCTGGTACCAAGTATCTCAGA	2015
1983	Qy	 GCCTCCGCTCTCTATCAGTCCGTATGCTTCAGAAAAGCTGTGGAAAGCTACTCTTCTCTTT	2042
2016	Db	 GCCTCCGCTCTCTATCAGTCCGTATGCTTCAGAAAAGCTGTGGAAAGCTACTCTTCTCTTT	2075
2043	Qy	 GAGACCTTGTGGAGAGGGCTGTGAAATGGTAGGCCAAAGAGAAATAGTTTCCCAGAGATAA	2102
2076	Db	 GAGACCTTGTGGAGAGGGCTGTGAAATGGTAGGCCAAAGAGAAATAGTTTCCCAGAGATAA	2135
2103	Qy	 AAACTGGTTGTTGGCCATGGCAGCCAAACGGAGGCTGAGAAATCCATCTCCACGAAGTCC	2162
2136	Db	 AAACTGGTTGTTGGCCATGGCAGCCAAACGGAGGCTGAGAAATCCATCTCCACGAAGTCC	2195
2163	Qy	 GTCAATCCAGACACCCAAATTCAGAGACAGAGCGGAAAGACATTTGCCAAGCCCGTCCAC	2222
2196	Db	 GTCAATCCAGACACCCAAATTCAGAGACAGAGCGGAAAGACATTTGCCAAGCCCGTCCAC	2255
2223	Qy	 CATCAAGCCCGAGCTTCCATGAGGAAAAATCTGCACATCTTCCATAGAAAGTCCCAGGAGGA	2282

D	b	2256	CATCAGCGCCGCTTCATTGAGGAAAAA	TCTGCACATACTTCCATAGAAAGTCCCAGGAGGA	23115
Q	y	2283	CTTCTGTGGTCCCTGAAACCACTCAA	CAGAGAATATAGATTCTTAATCTGAGTGAGTTACTGAGC	2342
D	b	2316	CTTCTGTGGTCTCTGAACTCACTCA	CAGAGAATATAGATTCTTAATCTGAGTGAGTTACTGAGC	2375
Q	y	2343	TTTGGTCCACTAAAACAAGCTGAGCT	TTGGTCCACTAAAACAAGATGAAAATAACAAG	2402
D	b	2376	TTTGGTCCACTAAAACAAGCTGAGCT	TTGGTCCACTAAAACAAGATGAAAATAACAAG	2435
Q	y	2403	TGACTCTATAACTCTGTGGTCTTTA	GAAAGCTGCCCTTTTATTTAGACAAAATCTTTTC	2462
D	b	2436	TGACTCTATAACTCTGTGGTCTTTA	GAAAGCTGCCCTTTTATTTAGACAAAATCTTTTC	2495
Q	y	2463	AACGCTGAAATGTACCTAAATCTGG	TCTCTACTACCATAATGTATATGCAGCTTCCCGAGGA	2522
D	b	2496	AACGCTGAAATGTACCTAAATCTGG	TCTCTACTACCATAATGTATATGCAGCTTCCCGAGGA	2555
Q	y	2523	TGAATGCTGTGTTTAAATTTTCAT	AAAGTAATTTGTGCATCTTAGCATTTTGAATGAATAG	2582
D	b	2556	TGAATGCTGTGTTTAAATTTTCAT	AAAGTAATTTGTGCATCTTAGCATTTTGAATGAATAG	2615
Q	y	2583	TCTTCACCTTTTTTAAATTTATTT	ATTTATTTTCTCTATATAATGACATCCCAGTTTCATGGAGCA	2642
D	b	2616	TCTTCACCTTTTTTAAATTTATTT	ATTTATTTTCTCTATATAATGACATCCCAGTTTCATGGAGCA	2675
Q	y	2643	AAAAACAAGTTTCTTGTGTTATCT	CGAAACTTTTCTATGCTCAGTGGAAGATATCTGCCAGCC	2702
D	b	2676	AAAAACAAGTTTCTTGTGTTATCT	CGAAACTTTTCTATGCTCAGTGGAAGATATCTGCCAGCC	2735
Q	y	2703	ACAGCATGAGCCCTGTGAAAGCTG	ACATGAGAAATCCCTCTGCTGAAGACCCCTGGTTCTGT	2762
D	b	2736	ACAGCATGAGCCCTGTGAAAGCTG	ACATGAGAAATCCCTCTGCTGAAGACCCCTGGTTCTGT	2795
Q	y	2763	TCTGCCCTCCAACATGATATATTT	TATTTTGAATACATAATCTTTTCACTA	2812
D	b	2796	TCTGCCCTCCAACATGATATATTT	TATTTTGAATACATAATCTTTTCACTA	2845
 RESULT 8 AAD63212 standard; DNA; 4221 BP.					
X	x	AC	AAD63212;		
X	x	DT	12-FEB-2004 (first entry)		
X	x	DE	Human DNA #37 used in the method for diagnosing cancer.		
X	x	KW	Human; breast cancer; metastasis; differential modulation; therapy; ds.		
X	s	OS	Homo sapiens.		
X	x	PN	US2003190656-A1.		
X	x	PD	09-OCT-2003.		
X	f	PF	21-MAR-2003; 2003US-00393590.		
X	x	PR	29-MAR-2002; 2002US-0368789P.		
X	a	PA	(WANG/) WANG Y.		
X	i	PI	Wang Y;		
X	r	DR	WPI; 2003-831621/77.		
X	t	PT	Prognosticating metastasis in a breast cancer patient comprises identifying differential modulation of each gene relative to the expression of the same genes in a normal population in combination of genes.		
X	s	PS	Claim 1; Page 122-124; Opp; English.		

XX The present invention relates to a method of prognosticating metastasis
CC in a breast cancer patient involves identifying differential modulation
CC of each gene relative to the expression of the same genes in a normal
CC population in combination of genes. The invention is useful for
CC prognosticating breast cancer in a patient. The present sequence is human
CC DNA used in the method for diagnosing cancer
XX

SQ Sequence 4221 BP; 1229 A; 897 C; 846 G; 1249 T; 0 U; 0 Other;

Query Match 98.5%; Score 2789.8; DB 10; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 22 CGATAACGATTGTCTGTGAGAGCGCAAGCGTGGCATTTCTGCTGAACTTGGAGGCATT 81
DB |||||
QY 82 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCGACCCCTGATGCTCTTCAATTCGGTG 141
DB |||||
QY 61 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCGACCCCTGATGCTCTTCAATTCGGTG 120
DB |||||
QY 142 CTCGCCAGCCAGCTTGGCGTCTCGAGAAATGATGGTCTTCAATACCCCTCTTCAA 201
DB |||||
QY 121 CTCGCCAGCCAGCTTGGCGTCTCGAGAAATGATGGTCTTCAATACCCCTCTTCAA 180
DB |||||
QY 202 TCCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACACACTTCTTATGGAGAAC 261
DB |||||
QY 181 TCCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACACACTTCTTATGGAGAAC 240
DB |||||
QY 262 GGAGTCCAGTTCCTCTTTGGATGTACCTTCTCTGCTGCCAATATGGACATGTA 321
DB |||||
QY 241 GGAGTCCAGTTCCTCTTTGGATGTACCTTCTCTGCTGCCAATATGGACATGTA 300
DB |||||
QY 322 CTAGCAGTTGCCAATGAAGAGCTTTGTTCGATTGTATTAACACAGAAATCAAAAGTTTC 381
DB |||||
QY 301 CTAGCAGTTGCCAATGAAGAGCTTTGTTCGATTGTATTAACACAGAAATCAAAAGTTTC 360
DB |||||
QY 382 AGAAAGAGTGTCTCAAGAAATGATGGTCTCACTGGAATCGCGTCTTTGACCTGGCCTGG 441
DB |||||
QY 361 AGAAAGAGTGTCTCAAGAAATGATGGTCTCACTGGAATCGCGTCTTTGACCTGGCCTGG 420
DB |||||
QY 442 GTTCTCTGTGAACTTAACTTTTACAGCAGCAGTGTATCAACAGCCCAAAATTTTGGAC 501
DB |||||
QY 421 GTTCTCTGTGAACTTAACTTTTACAGCAGCAGTGTATCAACAGCCCAAAATTTTGGAC 480
DB |||||
QY 502 GTAAAGCTGGTGAGTGTGAAATGCAAGGTCAATGAGCCTCAAGTCAGTT 561
DB |||||
QY 481 GTAAAGCTGGTGAGTGTGAAATGCAAGGTCAATGAGCCTCAAGTCAGTT 540
DB |||||
QY 562 GCCTTTTCTAAGTTTGAGAAAGCTGTATTTCTGTACGGGTGGAAGAGATGGCAACATTATG 621
DB |||||
QY 541 GCCTTTTCTAAGTTTGAGAAAGCTGTATTTCTGTACGGGTGGAAGAGATGGCAACATTATG 600
DB |||||
QY 622 GTCTGGGTACAGGTGCNACAAAAGATGGTTTTATAGCAGTGAATCAATCAGT 681
DB |||||
QY 601 GTCTGGGTACAGGTGCNACAAAAGATGGTTTTATAGCAGTGAATCAATCAGT 660
DB |||||
QY 682 GGAGTCTCAATACCTCAGACAAGCAACCCCTTCAAAACCCCAAGAAAGACAGAAATTC 741
DB |||||
QY 661 GGAGTCTCAATACCTCAGACAAGCAACCCCTTCAAAACCCCAAGAAAGACAGAAATTC 720
DB |||||
QY 742 AAAGGACTTGTCTCTGTGGATTTCAGCAAGAGTGTACTGTGGTCTCTCTTCAAGAC 801
DB |||||
QY 721 AAAGGACTTGTCTCTGTGGATTTCAGCAAGAGTGTACTGTGGTCTCTCTTCAAGAC 780
DB |||||
QY 802 GAGATACCTTAGTCTCAGCAGAGCTGTGGATGGGATTAATCAAGATATGGATTACGT 861
DB |||||
QY 781 GAGATACCTTAGTCTCAGCAGAGCTGTGGATGGGATTAATCAAGATATGGATTACGT 840
DB |||||
QY 862 AAGAAATTACTGCTTATCGACAAGAACCCATAGCATCCAAAGTCTTTCTGTACCCAGGT 921
DB |||||
QY 841 AAGAAATTACTGCTTATCGACAAGAACCCATAGCATCCAAAGTCTTTCTGTACCCAGGT 900
DB |||||

QY 922 AGCAGCACTCGAAAACTTGGATATTCAAGTCTGATTTTGGATTCCACTGGCTCTACTTTA 981
DB |||||
QY 901 AGCAGCACTCGAAAACTTGGATATTCAAGTCTGATTTTGGATTCCACTGGCTCTACTTTA 960
DB |||||
QY 982 TTTGCTAATTGCAAGAGATTAACATCTACATGTTTAAATATGACTGGTGTGAAGACTTCT 1041
DB |||||
QY 961 TTTGCTAATTGCAAGAGATTAACATCTACATGTTTAAATATGACTGGTGTGAAGACTTCT 1020
DB |||||
QY 1042 CCAGTGGCTATTCTTCAATGGACACCAGAACTCTACCTTTTATGTAATAATCCAGCCTTAGT 1101
DB |||||
QY 1021 CCAGTGGCTATTCTTCAATGGACACCAGAACTCTACCTTTTATGTAATAATCCAGCCTTAGT 1080
DB |||||
QY 1102 CCAGATGACCAAGTTTTTATGTCAGTGGCTCAAGTCAAGCTGCTCATATATGAAGGTC 1161
DB |||||
QY 1081 CCAGATGACCAAGTTTTTATGTCAGTGGCTCAAGTCAAGCTGCTCATATATGAAGGTC 1140
DB |||||
QY 1162 TCCACACCTGGCAACCTCTCTACTCTCTCTGGGTCAATCTCAAGAGGTCAAGTCTGTG 1221
DB |||||
QY 1141 TCCACACCTGGCAACCTCTCTACTCTCTCTGGGTCAATCTCAAGAGGTCAAGTCTGTG 1200
DB |||||
QY 1222 TGCTGGTGTCCATCTGACTTTCACAAAGATTGCTTACCTGTTCTGATGACAAATACACTAAAA 1281
DB |||||
QY 1201 TGCTGGTGTCCATCTGACTTTCACAAAGATTGCTTACCTGTTCTGATGACAAATACACTAAAA 1260
DB |||||
QY 1282 ATCTGGCGCTTGAATAGAGCTTAGAGGAAACACAGAGGTGATTAACCTTCCACCGTG 1341
DB |||||
QY 1261 ATCTGGCGCTTGAATAGAGCTTAGAGGAAACACAGAGGTGATTAACCTTCCACCGTG 1320
DB |||||
QY 1342 GTTTGGCGCTCTCAGAAAGAAAAAGAGTCAAGACCTGGCCTAGTAAACAGATAACGAGTAGC 1401
DB |||||
QY 1321 GTTTGGCGCTCTCAGAAAGAAAAAGAGTCAAGACCTGGCCTAGTAAACAGATAACGAGTAGC 1380
DB |||||
QY 1402 CAGAGTACTCTCTGCAAAAGCCCCCAGAGGTAAAGTGCAATCCATCCAATCTTCCCGGTCA 1461
DB |||||
QY 1381 CAGAGTACTCTCTGCAAAAGCCCCCAGAGGTAAAGTGCAATCCATCCAATCTTCCCGGTCA 1440
DB |||||
QY 1462 TCCGACGCTGTGCCCCCAGAGTGTCTGGAGACCTCTCCCTCTTCCATACACTCTCTACG 1521
DB |||||
QY 1441 TCCGACGCTGTGCCCCCAGAGTGTCTGGAGACCTCTCCCTCTTCCATACACTCTCTACG 1500
DB |||||
QY 1522 TTCTCTATTTAAAGCTCTCTCTGCAAGCCCGGTCTCCCATCAACAGAGAGGTCTGTCT 1581
DB |||||
QY 1501 TTCTCTATTTAAAGCTCTCTCTGCAAGCCCGGTCTCCCATCAACAGAGAGGTCTGTCT 1560
DB |||||
QY 1582 TCCTCCGTCTCTCCCAAGCCACCTTCTCATCTTTCAAGATGTGATTAGAAACTGGGTGACC 1641
DB |||||
QY 1561 TCCTCCGTCTCTCCCAAGCCACCTTCTCATCTTTCAAGATGTGATTAGAAACTGGGTGACC 1620
DB |||||
QY 1642 CGAACACTTCTCTCATCACCAACCCATCACTCCACCTGTTCCGGAGACCAAGATCATGTCT 1701
DB |||||
QY 1621 CGAACACTTCTCTCATCACCAACCCATCACTCCACCTGTTCCGGAGACCAAGATCATGTCT 1680
DB |||||
QY 1702 CCGAGAAAAAGCCCTTATTTCTGTGAGCCAGAAAGTCAATCCCAAGCAGAGGCTTGTCTGAG 1761
DB |||||
QY 1681 CCGAGAAAAAGCCCTTATTTCTGTGAGCCAGAAAGTCAATCCCAAGCAGAGGCTTGTCTGAG 1740
DB |||||
QY 1762 TCTAGAAATAGAGTAAAGAGAGGCTAGACTCAAGTGTCTGAGAGGTGTGAAACAAAG 1821
DB |||||
QY 1741 TCTAGAAATAGAGTAAAGAGAGGCTAGACTCAAGTGTCTGAGAGGTGTGAAACAAAG 1800
DB |||||
QY 1822 TGTGTGAAGAGTGTGTAAGTGTGACTGAGCTTCAATGAGTGTGAAATCTTCAATTTG 1881
DB |||||
QY 1801 TGTGTGAAGAGTGTGTAAGTGTGACTGAGCTTCAATGAGTGTGAAATCTTCAATTTG 1860
DB |||||
QY 1882 GATCTGTGCTGCCCTTGTCTGGTAAACAGGAAGACCTTAGTAGAGTCTCTAGGCTCTTACC 1941
DB |||||
QY 1861 GATCTGTGCTGCCCTTGTCTGGTAAACAGGAAGACCTTAGTAGAGTCTCTAGGCTCTTACC 1920
DB |||||
QY 1942 AAATCAAGCAAAATTTGAAGGAGCTGGTACAGTATCTCAGAGCCTCGGTCTCTTATCAGT 2001
DB |||||
QY 1921 AAATCAAGCAAAATTTGAAGGAGCTGGTACAGTATCTCAGAGCCTCGGTCTCTTATCAGT 1980
DB |||||

QY 2002 CCGTATGCTTTCAGAAAGCTGTGGAAGCTACCTCTTCTTTGAGACCTTGTGGAAGGG 2061
DB 1981 CCGTATGCTTTCAGAAAGCTGTGGAAGCTACCTCTTCTTTGAGACCTTGTGGAAGGG 2040
QY 2062 TCTGAAATGGTAGGCAAGAGAAATAGTTCCTCCAGAGAAATAAAAAAGTGTGTCCTCATG 2121
DB 2041 TCTGAAATGGTAGGCAAGAGAAATAGTTCCTCCAGAGAAATAAAAAAGTGTGTCCTCATG 2100
QY 2122 GCAGCCAAACGAGGCTGAGAAATCCATCTCCAGAGAAATCCGCTCATCCAGAGACCCCAAT 2181
DB 2101 GCAGCCAAACGAGGCTGAGAAATCCATCTCCAGAGAAATCCGCTCATCCAGAGACCCCAAT 2160
QY 2182 TCCAGAGACAGAGCGGAAAGCAATTCGCAAGCCGCTCACCATCAGCCGAGCTCCCATG 2241
DB 2161 TCCAGAGACAGAGCGGAAAGCAATTCGCAAGCCGCTCACCATCAGCCGAGCTCCCATG 2220
QY 2242 AGGAAATCTGCATATACATCTCCATAGAAAGTCCCAAGGAGCACTTCGTGGTCTCTGAAAC 2301
DB 2221 AGGAAATCTGCATATACATCTCCATAGAAAGTCCCAAGGAGCACTTCGTGGTCTCTGAAAC 2280
QY 2302 TCAACAGAAATATAGATCTTCTAGTGTAGTGTACTGAGCTTGGTCCACTTAAACACAG 2361
DB 2281 TCAACAGAAATATAGATCTTCTAGTGTAGTGTACTGAGCTTGGTCCACTTAAACACAG 2340
QY 2362 CTGAGCTTTGGTCCACTTAAACAGATGAAATACAGAGTACTCTATATCTCTGGTC 2421
DB 2341 CTGAGCTTTGGTCCACTTAAACAGATGAAATACAGAGTACTCTATATCTCTGGTC 2400
QY 2422 TTTAAGAAAGCTGCCCTTTTCATTTTTCAGCAAAATCTTTTCAACGCTGAAATGTACCTAA 2481
DB 2401 TTTAAGAAAGCTGCCCTTTTCATTTTTCAGCAAAATCTTTTCAACGCTGAAATGTACCTAA 2460
QY 2482 TCTGGTCTACATACATATATATATGAGCTTCCGAGGATGAATGCTGTGTTTAAAT 2541
DB 2461 TCTGGTCTACATACATATATATATGAGCTTCCGAGGATGAATGCTGTGTTTAAAT 2520
QY 2542 TCATAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATAT 2601
DB 2521 TCATAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATAT 2580
QY 2602 TCATCTTCTCTATAATATGACATCCAGTTCATGAGGCAAAACAAAGTTTCTTGTGA 2661
DB 2581 TCATCTTCTCTATAATATGACATCCAGTTCATGAGGCAAAACAAAGTTTCTTGTGA 2640
QY 2662 TCTGAAACTTTCTATGCTCAGTGAAGATATCTGCGAGCCACAGATGAGCCCTGTGA 2721
DB 2641 TCTGAAACTTTCTATGCTCAGTGAAGATATCTGCGAGCCACAGATGAGCCCTGTGA 2700
QY 2722 GGCTGACTGAGAAATCTCTGCTGAAGACCCCTGTTCTGTTCTGCTCCACATGTATA 2781
DB 2701 GGCTGACTGAGAAATCTCTGCTGAAGACCCCTGTTCTGTTCTGCTCCACATGTATA 2760
QY 2782 ATTTTATTGAAATACATAATCTTTTCACTATG 2814
DB 2761 ATTTTATTGAAATACATAATCTTTTCACTATG 2793

RESULT 9
ID AAD62813
AC AAD62813 standard; DNA; 4221 BP.
XX AAD62813;
XX
DT 12-FEB-2004 (first entry)
DE Human DNA #26 used in the method for diagnosing cancer.
XX Human; cancer; differential modulation; gene expression profile; ds.
XX Homo sapiens.
XX
PN US2003194733-A1.
XX

PD 16-OCT-2003.
XX 21-MAR-2003; 2003US-00393567.
XX 29-MAR-2002; 2002US-0368667P.
PA (WANG/) WANG Y.
PI Wang Y;
XX WPI; 2003-844450/78.
XX
PT Diagnosing cancer comprises identifying differential modulation of each
PT gene (relative to the expression of the same genes in a normal
PT population) in a combination of genes.
XX
PS Example 3; Page 122-124; Opp; English.
CC The invention relates to a method for diagnosing cancer. The method
CC comprising identifying differential modulation of each gene (relative to
CC the expression of the same genes in a normal population) in a combination
CC of genes. The method, diagnostic portfolio and the kit are useful in
CC diagnosing cancer. The method may also be used for determining gene
CC expression profiles. The present sequence is human DNA used to illustrate
CC the method of the invention
XX
SQ Sequence 4221 BP; 1229 A; 897 C; 846 G; 1249 T; 0 U; 0 Other;

Query Match 98.5%; Score 2789.8; DB 10; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 22 CGATAACGATTTGTGTTGTGAGAGCGCAACGTCGGATTTCTGCTGAACCTTGGAGGCATT 81
DB 1 CGATAACGATTTGTGTTGTGAGAGCGCAACGTCGGATTTCTGCTGAACCTTGGAGGCATT 60
QY 82 TCTAGACCTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG 141
DB 61 TCTAGACCTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG 120
QY 142 CTCGCGCAGCCCGAGCTTGGGCTCTGAGAAATGAGTGGTCTTCAATACCCCTTTCAA 201
DB 121 CTCGCGCAGCCCGAGCTTGGGCTCTGAGAAATGAGTGGTCTTCAATACCCCTTTCAA 180
QY 202 TCCCTTCTGACTGGTATCAGTGCAGTGTATGATGAACACACTTCTTATGGAGAAC 261
DB 181 TCCCTTCTGACTGGTATCAGTGCAGTGTATGATGAACACACTTCTTATGGAGAAC 240
QY 262 GGAGTCCAGTTCCTCCCTTTTGGATGTACCTTCTCTGCTCCCAATATGGAACATGTA 321
DB 241 GGAGTCCAGTTCCTCCCTTTTGGATGTACCTTCTCTGCTCCCAATATGGAACATGTA 300
QY 322 CTAGCAGTTGCCAATGAAAGAGGCTTTGTTTCGATTTGTATATAACAGAAATCAAAAGTTTC 381
DB 301 CTAGCAGTTGCCAATGAAAGAGGCTTTGTTTCGATTTGTATATAACAGAAATCAAAAGTTTC 360
QY 382 AGAAGAGTCTTCAAGAAATGAGTGCCTCAGTGGATGCGTCTTTCGCTGCGCTGG 441
DB 361 AGAAGAGTCTTCAAGAAATGAGTGCCTCAGTGGATGCGTCTTTCGCTGCGCTGG 420
QY 442 GTTCTGCTGAACTTAAACTTTTACAGCAGCAGTGTATCAACAGCCAAATTTTGGGAC 501
DB 421 GTTCTGCTGAACTTAAACTTTTACAGCAGCAGTGTATCAACAGCCAAATTTTGGGAC 480
QY 502 GTAAAAAGCTGGTGCCTGATTTGGAAACATGCAAGGTCTATCAATGCAGCTCAAGTCAGTT 561
DB 481 GTAAAAAGCTGGTGCCTGATTTGGAAACATGCAAGGTCTATCAATGCAGCTCAAGTCAGTT 540
QY 562 GCCTTTTCTAAAGTTTGAGAAAGCTGTATTTCTGACGGGTGGAAGATGGCAACATTATG 621
DB 541 GCCTTTTCTAAAGTTTGAGAAAGCTGTATTTCTGACGGGTGGAAGATGGCAACATTATG 600
QY 622 GTCTGGGATACAGGTGCAACAAAAAGATGGGTTTTTATAGGCAAGTGAATCAAAATCAGT 681

|||||
601 GTCTGGGATACGAGTGCAACAAAAAGATGGGTTTTATAGGCAAGTGAATCAATCAGT 660
QY
682 GGAGCTCAAAATACCTCAGACAAGCAAAACCCCTTCAAAACCCCAAGAGAAAACAGAAATTC 741
Db
661 GGAGCTCAAAATACCTCAGACAAGCAAAACCCCTTCAAAACCCCAAGAGAAAACAGAAATTC 720
QY
742 AAAGGACTTGCTCCTTCTGTGGATTTCCAGCMAAGTGTTACTGGGTCTCTTTCAAGAC 801
Db
721 AAAGGACTTGCTCCTTCTGTGGATTTCCAGCMAAGTGTTACTGGGTCTCTTTCAAGAC 780
QY
802 GAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAAGTATGGGATTAAGT 861
Db
781 GAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAAGTATGGGATTAAGT 840
QY
862 AAGAAATTAATCTGCTTATCGACAAGACCCATAGCATCCAAGTCTTTCTCTGTACCCAGGT 921
Db
841 AAGAAATTAATCTGCTTATCGACAAGACCCATAGCATCCAAGTCTTTCTCTGTACCCAGGT 900
QY
922 AGCAGACTCGAAATCTTGGATATTTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 981
Db
901 AGCAGACTCGAAATCTTGGATATTTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 960
QY
982 TTTTGCTAATTCACAGACGATAACATCTACATGTTTAAATATGACTGGGTTCGAAGACTTCT 1041
Db
961 TTTTGCTAATTCACAGACGATNAACATCTACATGTTTAAATATGACTGGGTTCGAAGACTTCT 1020
QY
1042 CCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTATGTAAATCCAGCCTTAGT 1101
Db
1021 CCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTATGTAAATCCAGCCTTAGT 1080
QY
1102 CCAGATGACAGTTTTATAGTCAAGTGGCTCAAGTGAAGCTGCCTACATATGGAGGTC 1161
Db
1081 CCAGATGACAGTTTTATAGTCAAGTGGCTCAAGTGAAGCTGCCTACATATGGAGGTC 1140
QY
1162 TCCACACCTCGCAACCTCTACTGCTCTCTGGGTCAATCTCAAGAGCTCAGCTCTGTG 1221
Db
1141 TCCACACCTCGCAACCTCTACTGCTCTCTGGGTCAATCTCAAGAGGTCAAGTCTGTG 1200
QY
1222 TGCTGGTGTCCATCTGACTTCAAAAGATGCTACCTGTTCTGTATGACAATACACTAAAA 1281
Db
1201 TGCTGGTGTCCATCTGACTTCAAAAGATGCTACCTGTTCTGTATGACAATACACTAAAA 1260
QY
1282 ATCTGGCGCTTGAATAGAGGCTTAGAGAGAAACAGAGAGTGATAAACTTTTCCAGGTC 1341
Db
1261 ATCTGGCGCTTGAATAGAGGCTTAGAGAGAAACAGAGAGTGATAAACTTTTCCAGGTC 1320
QY
1342 GGTTCGGGCTCTCAGAGAAAGAAAGAGTCAAGACCTGGCCTAGTAAACAGTAAACGAGTAGC 1401
Db
1321 GGTTCGGGCTCTCAGAGAAAGAAAGAGTCAAGACCTGGCCTAGTAAACAGTAAACGAGTAGC 1380
QY
1402 CAGAGTACTCTCTGCAAGAGGCTTCAAGTGAATCCATCCAATTTCTTCCCGGTCA 1461
Db
1381 CAGAGTACTCTCTGCAAGAGGCTTCAAGTGAATCCATCCAATTTCTTCCCGGTCA 1440
QY
1462 TCCGAGCTGTGCCCCAAGCTGCTGGAGACCTCCCTCTTCTTCAAAATCTCCTACG 1521
Db
1441 TCCGAGCTGTGCCCCAAGCTGCTGGAGACCTCCCTCTTCTTCAAAATCTCCTACG 1500
QY
1522 TTCTCTATTTAAACCTCTCTGCCAAGCGCGTCTCCCATCAACAGAGAGGCTCTGTC 1581
Db
1501 TTCTCTATTTAAACCTCTCTGCCAAGCGCGTCTCCCATCAACAGAGAGGCTCTGTC 1560
QY
1582 TCCTCCGCTCTCTCCCAAGCCACTTTCATCTTTCAAGATGTGATTTAGAACTGGGTGACC 1641
Db
1642 CGAACACCTTCTCTATCACCCCATCTCACTCCACCTGCTTGGAGACCAAGATCATGTCT 1701
QY
1621 CGAACACCTTCTCTATCACCCCATCTCACTCCACCTGCTTGGAGACCAAGATCATGTCT 1680
Db
1702 CCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCAATCCCAAGCAGAGGCTTGCTCTGAG 1761
QY
|||||

Db
1681 CCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCAATCCCAAGCAGAGGCTTGCTCTGAG 1740
QY
1762 TCTAGAAATAGATGATAAGAGAGGCTTAGACTCAAGCTGTCTGGAGAGTGTGAAAACAAAG 1821
Db
1741 TCTAGAAATAGATGATAAGAGAGGCTTAGACTCAAGCTGTCTGGAGAGTGTGAAAACAAAG 1800
QY
1822 TGTGTGAAGAGTTTCTAACTGTGTGACTGAGCTTGCATGSCCAAGTTGAAAATCTTTCATTG 1881
Db
1801 TGTGTGAAGAGTTTCTAACTGTGTGACTGAGCTTGCATGSCCAAGTTGAAAATCTTTCATTG 1860
QY
1882 GATCTGTGCTGCTTGTGTTAAACAGGAAGACCTTAGTAGAGACTCTCTAGAGTCTTAC 1941
Db
1861 GATCTGTGCTGCTTGTGTTAAACAGGAAGACCTTAGTAGAGACTCTCTAGAGTCTTAC 1920
QY
1942 AAATCAAGCAAAATTTGAAGAGCTGGTACCGATATCTCAGAGCCTCGTCTCTATCAGT 2001
Db
1921 AAATCAAGCAAAATTTGAAGAGCTGGTACCGATATCTCAGAGCCTCGTCTCTATCAGT 1980
QY
2002 CCGTATCTTTCAGAAAGCTGTGGAACGCTACCTCTCTCTTTTGAGACCTTGTGGAGAGGG 2061
Db
1981 CCGTATCTTTCAGAAAGCTGTGGAACGCTACCTCTCTCTTTTGAGACCTTGTGGAGAGGG 2040
QY
2062 TCTCAAAATGCTAGGCAAGAGATAGTTCCCGCAGAGAAATAAAACTGGTGTGGCCCATG 2121
Db
2041 TCTCAAAATGCTAGGCAAGAGATAGTTCCCGCAGAGAAATAAAACTGGTGTGGCCCATG 2100
QY
2122 GCAGCCAAACGGAAGGCTGAGATTCATCTCCAGCAAGTCCGTCATCCAGACACCCCAT 2181
Db
2101 GCAGCCAAACGGAAGGCTGAGATTCATCTCCAGCAAGTCCGTCATCCAGACACCCCAT 2160
QY
2182 TCCAGGAGACAGAGCGGAAGAGACATGCCCAGAGCCGGTCAACATCAGCCCGAGTCCCATG 2241
Db
2161 TCCAGGAGACAGAGCGGAAGAGACATGCCCAGAGCCGGTCAACATCAGCCCGAGTCCCATG 2220
QY
2242 AGGAAATCTGCAATATCTTCCATAGAAAGTCCAGGAGGACTCTCTGTGGTCTTGAACAC 2301
Db
2221 AGGAAATCTGCAATATCTTCCATAGAAAGTCCAGGAGGACTCTCTGTGGTCTTGAACAC 2280
QY
2302 TCAACAGAAATATAGATCTTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTAAACAG 2361
Db
2281 TCAACAGAAATATAGATCTTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTAAACAG 2340
QY
2362 CTGAGCTTTGGTCCACTAAACAGAAATGAAAATGAAAAGTACTCTATAACTCTGGTC 2421
Db
2341 CTGAGCTTTGGTCCACTAAACAGAAATGAAAATGAAAAGTACTCTATAACTCTGGTC 2400
QY
2422 TTTAAGAAAGCTGCTCTTTTCAATTTTATAGACAAAATCTTTTCAACGCTGAAATGTACCTAA 2481
Db
2401 TTTAAGAAAGCTGCTCTTTTCAATTTTATAGACAAAATCTTTTCAACGCTGAAATGTACCTAA 2460
QY
2482 TCTGGTCTCTACTACATAATGATATGAGCTTCCCGAGGATGAATGCTGTGTTAAAT 2541
Db
2461 TCTGGTCTCTACTACATAATGATATGAGCTTCCCGAGGATGAATGCTGTGTTAAAT 2520
QY
2542 TCATAAGTAAATTTGTCACCTCTAGCAATTTTGAATGAATAGTCTTCACTTTTAAATTTAT 2601
Db
2521 TCATAAGTAAATTTGTCACCTCTAGCAATTTTGAATGAATAGTCTTCACTTTTAAATTTAT 2580
QY
2602 TCATCTCTCTATATATATGATCAATCCAGTTCATGGAGGCAAAAACAAAGTTCTTGTGTA 2661
Db
2581 TCATCTCTCTATATATATGATCAATCCAGTTCATGGAGGCAAAAACAAAGTTCTTGTGTA 2640
QY
2662 TCCTGAAACTTTTCTATGCTGAGTGAAGATATCTGCCAGCCACAGCATGAGGCTGTGAA 2721
Db
2641 TCCTGAAACTTTTCTATGCTGAGTGAAGATATCTGCCAGCCACAGCATGAGGCTGTGAA 2700
QY
2722 GGCTGACTGAGAAATCTCTGCTCAAGACCCCTGTTCTGTTCTGCTCCCAACATGTATA 2781
Db
2701 GGCTGACTGAGAAATCTCTGCTCAAGACCCCTGTTCTGTTCTGCTCCCAACATGTATA 2760
QY
2782 ATTTTATTTGAAATACATAATCTTTTCACTATG 2814
Db
2761 ATTTTATTTGAAATACATAATCTTTTCACTATG 2793

RESULT 10
ABX95498
ID ABX95498 standard; cDNA; 4221 BP.
AC
XX
AC ABX95498;
DT
DT 01-JUL-2003 (first entry)
XX
DE Human retinoic acid-regulated nuclear matrix-associated protein cDNA.
XX
KW Nucleic acid identification; T cell receptor; T cell; tumour antigen;
KW antigen presenting cell; HLA; nucleic acid library; ELISPOT assay;
KW enzyme-linked immunospot; T cell receptor binding epitope;
KW cancer associated antigen; ramp; human; polytope; gene; ss;
KW retinoic acid-regulated nuclear matrix-associated protein.
OS
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 103..2295
FT /*tag= a
FT /product= "ramp"
FT /note= "retinoic acid-regulated nuclear matrix-associated
protein"
XX
PN US2003003485-A1.
XX
PD 02-JAN-2003.
XX
PF 14-MAY-2002; 2002US-00145396.
XX
PR 15-MAY-2001; 2001US-0291125P.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI Uenaka A, Nakayama E;
XX
DR WPI; 2003-416593/39.
DR P-PSDB; AB009611.
XX
XX
PT Identifying nucleic acids encoding epitopes that bind T cell receptors on
PT T cells when presented by human leukocyte antigen by using an ELISPOT
PT assay that detects factors secreted by T cell in response to the receptor
PT binding.
XX
PS Disclosure; Page 33-35; 51pp; English.
XX
CC The invention describes a method of identifying nucleic acid encoding an
CC epitope that binds to a T cell receptor on T cell. The method involves
CC coculturing antigen presenting cells that contain a nucleic acid library
CC and express HLA molecules presenting epitopes, with a T cell having a T
CC cell receptor that binds the epitope so that the T cell receptor binds
CC the epitope encoded by the nucleic acid library. Detecting a factor
CC secreted by the T cell in response to T cell receptor binding, using
CC ELISPOT (enzyme-linked immunospot) assay, and correlating factor
CC secretion with presence of a nucleic acid encoding the epitope. The
CC method is useful for identifying a nucleic acid molecule encoding the
CC epitope that specifically binds to a T cell receptor on a T cell when
CC presented by an HLA molecule. A second method described in the invention
CC is useful for identifying antigens that specifically bind to a T cell
CC receptor on a T cell. The method is useful for identifying cancer
CC associated antigens. This sequence encodes the human retinoic acid-
CC regulated nuclear matrix-associated protein (ramp) peptides of which can
CC be combined with peptides of one or more other cancer-associated antigens
CC to form polytopes
XX
SQ Sequence 4221 BP; 1229 A; 897 C; 846 G; 1249 T; 0 U; 0 Other;
Query Match 98.5%; Score 2789.8; DB 10; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 22 CGATAACGATTTGTGTTGTGAGAGCGCAACGTCGATTTCTGTGCTGAACCTTTGGAGCAATT 81
DB 1 CGATAACGATTTGTGTTGTGAGAGCGCAACGTCGATTTCTGTGCTGAACCTTTGGAGCAATT 60
QY 82 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG 141
DB 61 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG 120
QY 142 CTCGCCAGCCCGAGCTTGGCGTCTTGAGAAATGGATGGTCTTCAATACCTCTTCAA 201
DB 121 CTCGCCAGCCCGAGCTTGGCGTCTTGAGAAATGGATGGTCTTCAATACCTCTTCAA 180
QY 202 TCCCTTCTGACTGGTATATCAGTCAGTGGTAAATGATGAACACACTTCTTATGGAGAAACA 261
DB 181 TCCCTTCTGACTGGTATATCAGTCAGTGGTAAATGATGAACACACTTCTTATGGAGAAACA 240
QY 262 GGAGTCCCAGTTCCTCTCTTTTGGATGTACCTTCTCTCTGCTCCCAATATGGAACATGTA 321
DB 241 GGAGTCCCAGTTCCTCTCTTTTGGATGTACCTTCTCTCTGCTCCCAATATGGAACATGTA 300
QY 322 CTAGCAGTTGCCAATGAAGAAGGCTTTTGTTCGATTTGATATACACAGAAATCAAAAGTTTC 381
DB 301 CTAGCAGTTGCCAATGAAGAAGGCTTTTGTTCGATTTGATATACACAGAAATCAAAAGTTTC 360
QY 382 AGAAGAAGTGTCTTCAAGAATGATGGCTCACTGGAATGCCGTCTTTGACCTGGCCTGG 441
DB 361 AGAAGAAGTGTCTTCAAGAATGATGGCTCACTGGAATGCCGTCTTTGACCTGGCCTGG 420
QY 442 GTTCTCGTGAACTTAAACCTTTTACAGCAGCAGGTGATCAAAACAGCCAAATTTTGGGAC 501
DB 421 GTTCTCGTGAACTTAAACCTTTTACAGCAGCAGGTGATCAAAACAGCCAAATTTTGGGAC 480
QY 502 GTAAAGCTGTGAGCTGATTTGGAACATGCAAAAGTCAATCAATGCAGCCTCAAGTCAGTT 561
DB 481 GTAAAGCTGTGAGCTGATTTGGAACATGCAAAAGTCAATCAATGCAGCCTCAAGTCAGTT 540
QY 562 GCCTTTTCTTAAGTTTGAGAAAGCTGTATTCTGTACGGGTGGAAGATGGCAACATTATG 621
DB 541 GCCTTTTCTTAAGTTTGAGAAAGCTGTATTCTGTACGGGTGGAAGATGGCAACATTATG 600
QY 622 GTCTGGGATACAGGTGCAACAAAAAGATGGGTTTTATAGGCAAGTGAATCAAAATCAGT 681
DB 601 GTCTGGGATACAGGTGCAACAAAAAGATGGGTTTTATAGGCAAGTGAATCAAAATCAGT 660
QY 682 GGAGCTCAAAATACCTCAGACAGCAAAACCCCTTCAAAACCCCAAGAGAAACAGAAATCA 741
DB 661 GGAGCTCAAAATACCTCAGACAGCAAAACCCCTTCAAAACCCCAAGAGAAACAGAAATCA 720
QY 742 AAGGACTTGTCTCTCTGCTGATTTCCAGCAAAAGTGTACTGTGCTCTCTTCAAGAC 801
DB 721 AAGGACTTGTCTCTCTGCTGATTTCCAGCAAAAGTGTACTGTGCTCTCTTCAAGAC 780
QY 802 GAGAAATACCTTAGTCTCAGCAGGAGCTGTGATGGGTAATCAAAATATGGATTTAGCT 861
DB 781 GAGAAATACCTTAGTCTCAGCAGGAGCTGTGATGGGTAATCAAAATATGGATTTAGCT 840
QY 862 AAGAAATATACCTTATTCGACAGAACCCATAGATCCAAAGTCTTCTGTACCCAGT 921
DB 841 AAGAAATATACCTTATTCGACAGAACCCATAGATCCAAAGTCTTCTGTACCCAGT 900
QY 922 AGCAGCACTCGAAAACCTTGGATATTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 981
DB 901 AGCAGCACTCGAAAACCTTGGATATTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 960
QY 982 TTTGCTAATTCACAGACGATAAATCTACATGTTTAAATATGATGGGTGGAAGACTTCT 1041
DB 961 TTTGCTAATTCACAGACGATAAATCTACATGTTTAAATATGATGGGTGGAAGACTTCT 1020
QY 1042 CCAGTGGCTATTTCAATGGACACAGAACTCTACCTTTTATGTAATAATCCAGCCTTAGT 1101
DB 1021 CCAGTGGCTATTTCAATGGACACAGAACTCTACCTTTTATGTAATAATCCAGCCTTAGT 1080

Qy	1102	CCAGATGACCAAGTTTTTTAGTGCAGTGGCTCAAGTGATGAAGCTGCTACATATATGGAAGGTC	11161
Db	1081	CCAGATGACCAAGTTTTTTAGTGCAGTGGCTCAAGTGATGAAGCTGCTACATATGGAAGGTC	11140
Qy	1162	TCCACACCTGGCAACCTCTCTACTGTGTCTCTCTGGGTCAATTTCTCAAGAGGTCAGTCTGTG	1221
Db	1141	TCCACACCTGGCAACCTCTCTACTGTGTCTCTCTGGGTCAATTTCTCAAGAGGTCAGTCTGTG	1200
Qy	1222	TGCTGGTGTCCATCTGACTTTCACAAAGATTGCTACCTGTTTCTGATGACAATACACTAAAA	1281
Db	1201	TGCTGGTGTCCATCTGACTTTCACAAAGATTGCTACCTGTTTCTGATGACAATACACTAAAA	1260
Qy	1282	ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACACGAGGTTGATAAACTTTCCACGGTG	1341
Db	1261	ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACACGAGGTTGATAAACTTTCCACGGTG	1320
Qy	1342	GGTTGGGCGCTCTCAGAGAGAAAAAGAGTCAAGACCTGGCCCTAGTAAACAGTTAAACGATGAC	1401
Db	1321	GGTTGGGCGCTCTCAGAGAGAAAAAGAGTCAAGACCTGGCCCTAGTAAACAGTTAAACGATGAC	1380
Qy	1402	CAGAGTACTCTCTGCGCAAGCCCCCAGAGGTTAAGTGCAATCATCCCAATTTCTCCCGTGCA	1461
Db	1381	CAGAGTACTCTCTGCGCAAGCCCCCAGAGGTTAAGTGCAATCATCCCAATTTCTCCCGTGCA	1440
Qy	1462	TCCGACGCTGTGTGCCCAAGCTGTGTGGAGACTCCCTCTTCTTCAAAATATCTCTTAGC	1521
Db	1441	TCCGACGCTGTGTGCCCAAGCTGTGTGGAGACTCCCTCTTCTTCAAAATATCTCTTAGC	1500
Qy	1522	TTCTCTATTAAAACTCTCTCTGCCAAGCCCGGTCTCCATCAACAGAGAGGCTCTGTCT	1581
Db	1501	TTCTCTATTAAAACTCTCTCTGCCAAGCCCGGTCTCCATCAACAGAGAGGCTCTGTCT	1560
Qy	1582	TCCTCCGCTCTCTCCCAAGCCACCTTCATCTTTCAAGATGTCGATTAGAACTGGGTGACC	1641
Db	1561	TCCTCCGCTCTCTCCCAAGCCACCTTCATCTTTCAAGATGTCGATTAGAACTGGGTGACC	1620
Qy	1642	CGAACACCTTCTCTCATCACCAACCATCACTCCACTGCTTCGGAGACCAAGATCATGTCT	1701
Db	1621	CGAACACCTTCTCTCATCACCAACCATCACTCCACTGCTTCGGAGACCAAGATCATGTCT	1680
Qy	1702	CCGAGAAAAAGCCCTTATTTCTGTGAGCGCAGAGTCATCCCAAGCAGAGGCTTGCTCTGAG	1761
Db	1681	CCGAGAAAAAGCCCTTATTTCTGTGAGCGCAGAGTCATCCCAAGCAGAGGCTTGCTCTGAG	1740
Qy	1762	TCTAGAAATAGAGTAAAGAGAGGCTAGACTCAAGCTGTCTGGAGAGTGTTGAACAAAAAG	1821
Db	1741	TCTAGAAATAGAGTAAAGAGAGGCTAGACTCAAGCTGTCTGGAGAGTGTTGAACAAAAAG	1800
Qy	1822	TGTTGGAAGAGTTGTAACTGTGTGACTGAGCTTGATGGCCCAAGTTGAAAATCTTTCATTG	1881
Db	1801	TGTTGGAAGAGTTGTAACTGTGTGACTGAGCTTGATGGCCCAAGTTGAAAATCTTTCATTG	1860
Qy	1882	GATCTGTCTGCCCTTGTGTGTAAACAGAGAGACCTTAGTAGAGNCTCTCTAGGTCCTTACC	1941
Db	1861	GATCTGTCTGCCCTTGTGTGTGTAAACAGAGAGACCTTAGTAGAGNCTCTCTAGGTCCTTACC	1920
Qy	1942	AAATCAAGCAAAATTTGAAGGAGCTGGTACCAAGTATCTCAGAGCCCTCGTCTCTTATCAGT	2001
Db	1921	AAATCAAGCAAAATTTGAAGGAGCTGGTACCAAGTATCTCAGAGCCCTCGTCTCTTATCAGT	1980
Qy	2002	CCGTATGCTTCAGAAAAGCTGTGGAAACGCTACTCTCTCTTTTGAGACCTTGTGGAGAGGG	2061
Db	1981	CCGTATGCTTCAGAAAAGCTGTGGAAACGCTACTCTCTCTTTTGAGACCTTGTGGAGAGGG	2040
Qy	2062	TCTGAAATGGTAGGCAAGAGAAATAGTTTCCCAGAGAAATAAAACTGGTTGTTGGCCATG	2121
Db	2041	TCTGAAATGGTAGGCAAGAGAAATAGTTTCCCAGAGAAATAAAACTGGTTGTTGGCCATG	2100
Qy	2122	GCAGCCAAACGGAAGGCTGAGAAATCCATCTCCAGAGAGTCGCTCATCCAGAGACCCCAAT	2181
Db	2101	GCAGCCAAACGGAAGGCTGAGAAATCCATCTCCAGAGAGTCGCTCATCCAGAGACCCCAAT	2160
Qy	2182	TCCAGGAGACAGACGGAAGAGACATTCGCAAGCCCGGTCCACCATCAAGCCGAGCTCCATG	2241

Db	2161		TCCAGGAGACAGAGCGGAAGACATTCGCAAGCCCGGTCAACATCAGGCCAGCTCCATG	2220
Qy	2242		AGGAAAATCTGCACATACTTCCATAGAAAAGTCCGAGGAGACTTCTGTGTCCTGAACAC	2301
Db	2221		AGGAAAATCTGCACATACTTCCATAGAAAAGTCCGAGGAGACTTCTGTGTCCTGAACAC	2280
Qy	2302		TCAACAGAAATATAGATTCTTAATCTGAGTGAGTTACTGAGCTTTGGTCCCACTAAACAAG	2361
Db	2281		TCAACAGAAATATAGATTCTTAATCTGAGTGAGTTACTGAGCTTTGGTCCCACTAAACAAG	2340
Qy	2362		CTGAGCTTTGGTCCACATAAACAAGATGAAAAATACAAGAGTGACTCTATAAATCTCTGGTC	2421
Db	2341		CTGAGCTTTGGTCCACTTAAACAAGATGAAAAATACAAGAGTGACTCTATAAATCTCTGGTC	2400
Qy	2422		TTTAAGAAAGCTGCCTTTTCATTTTATAGACAAAATCTTTTCAACGCTGAAAATGTACTTAA	2481
Db	2401		TTTAAGAAAGCTGCCTTTTCATTTTATAGACAAAATCTTTTCAACGCTGAAAATGTACTTAA	2460
Qy	2482		TCTGGTTCTACTACCAATAATGATATGCAGCTTCCCGAGGATGAATGCTGTGTTTAAATT	2541
Db	2461		TCTGGTTCTACTACCAATAATGATATGCAGCTTCCCGAGGATGAATGCTGTGTTTAAATT	2520
Qy	2542		TCATAAAGTAAAATTTGTCACTCTAGCAATTTTGAATGAATAGTCTTCACTTTTAAATAT	2601
Db	2521		TCATAAAGTAAAATTTGTCACTCTAGCAATTTTGAATGAATAGTCTTCACTTTTAAATAT	2580
Qy	2602		TCATCTTCTCTATAAATACATACCCAGTTCATGAGGCAAAAACAAAGTTTCTTGTTA	2661
Db	2581		TCATCTTCTCTATAAATACATACCCAGTTCATGAGGCAAAAACAAAGTTTCTTGTTA	2640
Qy	2662		TCCTGAAACTTTCTATGCTCAGTGGAAGTATCTGCCAGGCCACGCAATGAGGCTGTGAA	2721
Db	2641		TCCTGAAACTTTCTATGCTCAGTGGAAGTATCTGCCAGGCCACGCAATGAGGCTGTGAA	2700
Qy	2722		GGCTGACTGAGAAATCTCTGCTGNAGACCCCTGGTTCTGTCTTCTGCTCCAAACATGTATA	2781
Db	2701		GGCTGACTGAGAAATCTCTGCTGNAGACCCCTGGTTCTGTCTTCTGCTCCAAACATGTATA	2760
Qy	2782		ATTTTATTTGAAAATACATAAATCTTTTTCACATATG	2814
Db	2761		ATTTTATTTGAAAATACATAAATCTTTTTCACATATG	2793

RESULT 11	
ADP88308	
ID	ADP88308 standard; cDNA; 4221 BP.
XX	
AC	ADP88308;
XX	
DT	12-AUG-2004 (first entry)
XX	
DE	Metastatic breast cancer marker gene RAMP, SEQ ID NO:85.
XX	
KW	Diagnostic marker; portfolio; diagnostic parameter; gene expression;
KW	relative expression; heuristic rule; molecular diagnostic; diagnosis;
KW	prognosis; cancer; breast; prostate; ovarian; colorectal; lung;
KW	prognostic portfolio; human; metastatic; RAMP; gene; ss.
XX	
OS	Homo sapiens.
XX	
PN	EPI349104-A2.
XX	
PD	01-OCT-2003.
XX	
PF	31-MAR-2003; 2003EP-00252027.
XX	
PR	29-MAR-2002; 2002US-0368790P.
XX	
PA	(ORTH) ORTHO CLINICAL DIAGNOSTICS INC.
XX	
PI	Jatkoe T;

RESULT 11
ADP88308
ID ADP88308 standard: cDNA: 4221 BP.

DR WPI: 2003-815043/77.
XX Selecting a portfolio of markers useful in diagnostic applications
PT comprises defining diagnostic parameters and establishing a relationship
PT among the diagnostic parameters that identifies an optimized portfolio of
XX markers.
XX
PS Example 3; SEQ ID NO 85; 154pp; English.
XX
CC The invention relates to a method of selecting a portfolio of markers
CC (e.g., genes) for use in a diagnostic application. The method involves:
CC (a) defining diagnostic parameters; (b) establishing a relationship among
CC the diagnostic parameters so that they are optimised; and (c) selecting
CC the optimal group of markers (the portfolio) for the diagnostic
CC application. The diagnostic parameters can include a measure of the
CC relative degree of expression of a gene, a measure of the variation in
CC the measurement of the degree of gene expression, and the relationship
CC between the diagnostic and discriminating parameters can be a mean
CC variance relationship. The method further comprises the application of a
CC heuristic rule. The invention also relates to a general-purpose computer
CC programmed to identify a portfolio of markers, instructions for
CC performing the method of the invention, and a diagnostic portfolio
CC comprising genes selected according to the method. The method of the
CC invention can be used in conjunction with any method for determining the
CC gene expression of relevant cells as well as protein-based methods of
CC determining gene expression. The method of the invention can be used in
CC screening, diagnosing and prognosing diseases associated with changes in
CC the expression of multiple genes (e.g., breast, prostate, ovarian,
CC colorectal or lung cancer). By grouping these markers into portfolios,
CC the most reliable results can be obtained with the smallest number of
CC markers necessary to obtain such a result, which is particularly
CC important in multiple-step assays such as nucleic amplification methods.
CC The present sequence represents a human diagnostic marker gene that is a
CC member of an optimised portfolio of 28 genes for prognosing breast
XX cancer.
XX
SQ Sequence 4221 BP; 1229 A; 897 C; 846 G; 1249 T; 0 U; 0 Other;

Query Match 98.5%; Score 2789.8; DB 11; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 22 CGATAACGATTGTGTGTGAGAGGCGCAACGTGCGATTTCTGTGTAACCTTGGAGGCATT 81
DB 1 CGATAACGATTGTGTGTGAGAGGCGCAACGTGCGATTTCTGTGTAACCTTGGAGGCATT 60
QY 82 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCCGGTG 141
DB 61 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCCGGTG 120
QY 142 CTCGGCAGCCCGAGCTTGGGCTCTGAGAAATGGATGGTCTTCAATACCTCTTCAA 201
DB 121 CTCGGCAGCCCGAGCTTGGGCTCTGAGAAATGGATGGTCTTCAATACCTCTTCAA 180
QY 202 TCCCTTCTGACGTGTTATCAGTGCAGTGGTATGATGACACACTTCTTATGAGAAACA 261
DB 181 TCCCTTCTGACGTGTTATCAGTGCAGTGGTATGATGACACACTTCTTATGAGAAACA 240
QY 262 GGAGTCCCAAGTTCCTCTTTGGATGATACCTTCTTCTGCTCCCAATATGAAACATGTA 321
DB 241 GGAGTCCCAAGTTCCTCTTTGGATGATACCTTCTTCTGCTCCCAATATGAAACATGTA 300
QY 322 CTAGCAGTTGCCAATGAGAGGCTTTGTTGATGATGATGATGATGATGATGATGATGATGAT 381
DB 301 CTAGCAGTTGCCAATGAGAGGCTTTGTTGATGATGATGATGATGATGATGATGATGATGAT 360
QY 382 AGAAGAAGTGTCTCAAGAATGGATGGCTCTACTGGATGGCTTTTGGACCTGGCTGG 441
DB 361 AGAAGAAGTGTCTCAAGAATGGATGGCTCTACTGGATGGCTTTTGGACCTGGCTGG 420
QY 442 GTTCTGTGGTAACTTAACTTGTGACAGCAGAGGCTGATCAACAGCCAAATTTTGGGAC 501
DB 421 GTTCTGTGGTAACTTAACTTGTGACAGCAGAGGCTGATCAACAGCCAAATTTTGGGAC 480

QY 502 GTAAAAGCTGTGAGCTGATTGGAAACATGCAAAAGCTCATCAATGCAGCCTCAAGTCACTT 561
DB 481 GTAAAAGCTGTGAGCTGATTGGAAACATGCAAAAGCTCATCAATGCAGCCTCAAGTCACTT 540
QY 562 GCCTTTTCTAAAGTTTGAGAAAGCTGTATTCTGTGCGGTGAAAGAGATGGCAACATTATG 621
DB 541 GCCTTTTCTAAAGTTTGAGAAAGCTGTATTCTGTGCGGTGAAAGAGATGGCAACATTATG 600
QY 622 GTCTGGGATACAGGTGCAACAAAAGATGGGTTTTATATGCGAGTCAATCAATCACTAGT 681
DB 601 GTCTGGGATACAGGTGCAACAAAAGATGGGTTTTATATGCGAGTCAATCAATCACTAGT 660
QY 682 GGAGCTCACAATACCTCAGACAAGCAACCCCTTCAAAACCCCAAGAAAGAAACAGAAATCA 741
DB 661 GGAGCTCACAATACCTCAGACAAGCAACCCCTTCAAAACCCCAAGAAAGAAACAGAAATCA 720
QY 742 AAAGGACTTGTCTCTCTGTGGATTTCCAGCAAAAGTGTATTCTGTGGTCTCTTTCAAGAC 801
DB 721 AAAGGACTTGTCTCTCTGTGGATTTCCAGCAAAAGTGTATTCTGTGGTCTCTTTCAAGAC 780
QY 802 GAGATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAGTATGGATTTAGT 861
DB 781 GAGATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAGTATGGATTTAGT 840
QY 862 AAGATATTACTGTCTTATCGACAAGAACCCATAGCATCCAAAGTCTTTCTGTACCAGCT 921
DB 841 AAGATATTACTGTCTTATCGACAAGAACCCATAGCATCCAAAGTCTTTCTGTACCAGCT 900
QY 922 AGCAGCACTCGAAAACCTTGGATATTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 981
DB 901 AGCAGCACTCGAAAACCTTGGATATTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 960
QY 982 TTTGCTAATTGCACAGACGATAATCTACATGTATTATATGATGACTGGTGAAGACTTCT 1041
DB 961 TTTGCTAATTGCACAGACGATAATCTACATGTATTATATGACTGGTGAAGACTTCT 1020
QY 1042 CCAGTGGCTATTTCAATGGACACCAAGAACTCTACTTTTATGTAATAATCCAGCCTTAGT 1101
DB 1021 CCAGTGGCTATTTCAATGGACACCAAGAACTCTACTTTTATGTAATAATCCAGCCTTAGT 1080
QY 1102 CCAGATGACCAAGTTTATGTCAGTGGCTCAAGTGAATGAGCTGCCTACATATGGAAGTCT 1161
DB 1081 CCAGATGACCAAGTTTATGTCAGTGGCTCAAGTGAATGAGCTGCCTACATATGGAAGTCT 1140
QY 1162 TCCACACCTGGCAACCTCTACTGTCTCTGGTCAATCTCAAGAGGTCACTCTGTG 1221
DB 1141 TCCACACCTGGCAACCTCTACTGTCTCTGGTCAATCTCAAGAGGTCACTCTGTG 1200
QY 1222 TGCTGGTGTCCATCTGACTTCCAAAAGATTCTCTCTGATGACAATACACTAAAA 1281
DB 1201 TGCTGGTGTCCATCTGACTTCCAAAAGATTCTCTCTGATGACAATACACTAAAA 1260
QY 1282 ATCTGGCCCTTTGAATAGAGGCTTAGAGGAGAAAACAGAGAGGTGATAAATTTTCCACGGTG 1341
DB 1261 ATCTGGCCCTTTGAATAGAGGCTTAGAGGAGAAAACAGAGAGGTGATAAATTTTCCACGGTG 1320
QY 1342 GGTGGGGCTCTCAGAGAAAAGAGTCAAGACCTGGCTTAGTAAACAGTAACGAGTAGC 1401
DB 1321 GGTGGGGCTCTCAGAGAAAAGAGTCAAGACCTGGCTTAGTAAACAGTAACGAGTAGC 1380
QY 1402 CAGATACTCTCTGCAAGAGCCCGAGGGTAAAGTGAATCCATCCAAATTTTCCCGCTCA 1461
DB 1381 CAGATACTCTCTGCAAGAGCCCGAGGGTAAAGTGAATCCATCCAAATTTTCCCGCTCA 1440
QY 1462 TCCGAGCTGTGCCCCCAAGCTGTGTCGAGACCTCTCTCTCTTCAAAATCTCTTACG 1521
DB 1441 TCCGAGCTGTGCCCCCAAGCTGTGTCGAGACCTCTCTCTTCAAAATCTCTTACG 1500
QY 1522 TTCTCTATTAAAACTCTCTCTGCCAAGCCCGGTCTCCCATCAACAGAAAGAGGTCTGTCT 1581
DB 1501 TTCTCTATTAAAACTCTCTCTGCCAAGCCCGGTCTCCCATCAACAGAAAGAGGTCTGTCT 1560

QY 1582 TCCTCGGTCTCTCCCAAGCCACCTTCATCTTTCAAGATGTCGATTAGAACTGGGTGACC 1641
DB 1561 TCCTCGGTCTCTCCCAAGCCACCTTCATCTTTCAAGATGTCGATTAGAACTGGGTGACC 1620
QY 1642 CGAAACACCTTCCTCATCACACCCATCACTCCACCTGCTTCGGAGACCAAGATCATGTCT 1701
DB 1621 CGAACACCTTCCTCATCACACCCATCACTCCACCTGCTTCGGAGACCAAGATCATGTCT 1680
QY 1702 CCGAGAAAAGCCCTTTATTCCTGTGAGCCAGAAAGTCAATCCCAAGCAGAGGCTTGCTCTGAG 1761
DB 1681 CCGAGAAAAGCCCTTTATTCCTGTGAGCCAGAAAGTCAATCCCAAGCAGAGGCTTGCTCTGAG 1740
QY 1762 TCTAGAAATAGAGTAAGAGGAGCTAGACTCAAGCTGCTGGAGAGTGTGAAACAAAAG 1821
DB 1741 TCTAGAAATAGAGTAAGAGGAGGCTAGACTCAAGCTGCTGGAGAGTGTGAAACAAAAG 1800
QY 1822 TGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGATGGCCAGTTGAAATCTTCATTTTG 1881
DB 1801 TGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGATGGCCAGTTGAAATCTTCATTTTG 1860
QY 1882 GATCTGTGCTGCTTCTGCTGTAACCAAGGAGACCTTAGTAAGGACTCTCTAGGTCCTTACC 1941
DB 1861 GATCTGTGCTGCTTCTGCTGTAACCAAGGAGACCTTAGTAAGGACTCTCTAGGTCCTTACC 1920
QY 1942 AAATCAAGCAAAATTTGAAGGAGCTGGTACCAGTATCTCAGAGCCTCCGTCCTCTATCAGT 2001
DB 1921 AAATCAAGCAAAATTTGAAGGAGCTGGTACCAGTATCTCAGAGCCTCCGTCCTCTATCAGT 1980
QY 2002 CCGTATGCTTCAGAAAGCTGTGGAAGCGTACCTCTTCTTGAGACCTTTGTGGAGAGGG 2061
DB 1981 CCGTATGCTTCAGAAAGCTGTGGAAGCGTACCTCTTCTTGAGACCTTTGTGGAGAGGG 2040
QY 2062 TCTGAAATGGTAGGCAAGAGAAATAGTTCCTCCAGAGAAATAAAACTGGTTGTGGCCCATG 2121
DB 2041 TCTGAAATGGTAGGCAAGAGAAATAGTTCCTCCAGAGAAATAAAACTGGTTGTGGCCCATG 2100
QY 2122 GCAGCCAAACGGAAGCTGAGAAATCAATCTCCAGAACTCCGTCATCCAGACACCCCAAT 2181
DB 2101 GCAGCCAAACGGAAGCTGAGAAATCAATCTCCAGAACTCCGTCATCCAGACACCCCAAT 2160
QY 2182 TCCAGGACACAGCGGAAAGACATTTGCCAAGCCCGGTCAACCATCAGCCAGCTCCCATG 2241
DB 2161 TCAGAGACACAGCGGAAAGACATTTGCCAAGCCCGGTCAACCATCAGCCAGCTCCCATG 2220
QY 2242 AGGAAATCTGCATACACTTCCATAGAAAGTCCCAAGGAGACTTCTGTGGTCCCTGAACAC 2301
DB 2221 AGGAAATCTGCATACACTTCCATAGAAAGTCCCAAGGAGACTTCTGTGGTCCCTGAACAC 2280
QY 2302 TCAACAGAAATATAGATTTCTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTTAAACAAAG 2361
DB 2281 TCAACAGAAATATAGATTTCTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTTAAACAAAG 2340
QY 2362 CTGAGCTTTGGTCCACTTAAACAGATGAAATACAGAGTGACTCTATTAATCTCTGGTC 2421
DB 2341 CTGAGCTTTGGTCCACTTAAACAGATGAAATACAGAGTGACTCTATTAATCTCTGGTC 2400
QY 2422 TTTAAGAAAGCTGCCCTTTTCATTTTATAGACAAATCTTTTCAACGCTGAAATGTACCTAA 2481
DB 2401 TTTAAGAAAGCTGCCCTTTTCATTTTATAGACAAATCTTTTCAACGCTGAAATGTACCTAA 2460
QY 2482 TCTGGTTCTACTACCAATATGATATGAGCTTCGGAGGATGAATGCTGTGTTTAAATTT 2541
DB 2461 TCTGGTTCTACTACCAATATGATATGAGCTTCGGAGGATGAATGCTGTGTTTAAATTT 2520
QY 2542 TCATAAAGTAAATTTGTCACTCTAGCATTTTCAATGAATGAGTCTTCACTTTTAAATTTAT 2601
DB 2521 TCATAAAGTAAATTTGTCACTCTAGCATTTTCAATGAATGAGTCTTCACTTTTAAATTTAT 2580
QY 2602 TCATCTTCTCTATAATAATGACATCCAGTTTCATGGAGGCAAAAAACAAGTTCTTGTTA 2661
DB 2581 TCATCTTCTCTATAATAATGACATCCAGTTTCATGGAGGCAAAAAACAAGTTCTTGTTA 2640
QY 2662 TCTGAAACCTTTCTATGCTCAGTGGGAAAGTATCTGCCAGCCACAGCATGAGCCCTGTGNA 2721

DB 2641 TCCTGAAACCTTTCTATGCTCAGTGGAAAGATATCTGCCAGCCACAGCATGAGCCCTGTGAA 2700
QY 2722 GGCTGACTGAGAAATCCTCTGCTGAAGACCCCTGGTTCTGTTCCTCCCAACATGTATA 2781
DB 2701 GGCTGACTGAGAAATCCTCTGCTGAAGACCCCTGGTTCTGTTCCTCCCAACATGTATA 2760
QY 2782 ATTTTATTTGAAATACATAATCTTTTCACTATG 2814
DB 2761 ATTTTATTTGAAATACATAATCTTTTCACTATG 2793

RESULT 12

ADP07299 standard; DNA; 4221 BP.

XX ADP07299;

XX 29-JUL-2004 (first entry)

XX Human L2DTL DNA.

XX ds; proliferative disease; breast; methylation; CpG; bisulfite; human.

XX Homo sapiens.

XX DE10255104-A1.

XX 11-MAR-2004.

XX 26-NOV-2002; 2002DE-01055104.

XX 27-AUG-2002; 2002DE-01039313.

XX (EPIG-) EPIGENOMICS AG.

XX Maier S;

XX WPI; 2004-284340/27.

XX GENBANK; NM_016448.

XX Analyzing proliferative diseases of breast cells, useful e.g. for diagnosis, prognosis and treatment of breast cancer, by determining methylation status of specific genes.

XX Claim 1; Page; 22pp; German.

XX This invention describes a novel method of analysing proliferative diseases of breast cells by determining the methylation status of certain genes. The invention also describes nucleic acids or their complements, oligomers, especially oligonucleotides or peptide nucleic acid oligomers, that hybridise to, or are identical with, any of the nucleic acids, the preparation of an oligomer array for analysing proliferative diseases of breast cells that are associated with the methylation status of CpG dinucleotides of the genes by bonding at least one oligomer to a solid phase and a kit comprising a bisulfite reagent and the oligomer. The nucleic acids are genomic sequences (5' - and/or regulatory and/or CpG-rich regions). The base sequence of the oligomer includes at least one CpG island, especially with C in the middle third of the sequence. The process involves a genomic DNA sample treated chemically, specifically with a bisulfite reagent, to convert non-methylated C to uracil or some other base having base-pairing properties different from those of C. Fragments of the treated DNA are then amplified, using primers and a polymerase and the methylation status of the genomic CpG dinucleotides is determined by analysis of the amplicons, particularly by hybridisation to the oligomer, optionally with extension of the hybridised oligomer by at least one base, or detection is by sequencing. The amplification may use methylation-specific primers. Alternatively, genomic DNA is extracted from a sample and digested with methylation-specific restriction enzymes, then the digestion fragments detected, optionally after amplification. In either method, more than 10 fragments of 100-200 bp are amplified in a single reaction vessel, using a heat-resistant DNA polymerase in PCR. The amplicons carry detectable markers, e.g. fluorophores, radioisotopes

CC and/or releasable fragments of known mass that can be detected by mass
CC spectrometry. The method is used for characterisation, classification,
CC differentiation, staging, phase-estimation, diagnosis and/or therapy of
CC proliferative diseases of breast cells. The method provides very specific
CC classification of proliferative diseases, allowing better treatment. It
CC can both characterise methylation status and detect single-nucleotide
CC polymorphisms. This sequence represents human gene used to illustrate the
CC method of the invention. NOTE: This sequence does not appear in the
CC printed specification but has been retrieved from Genbank.

XX SQ Sequence 4221 BP; 1229 A; 897 C; 846 G; 1249 T; 0 U; 0 Other;

Query Match 98.5%; Score 2789.8; DB 12; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	22	CGATAACGATTGTGTGTGAGAGCGCAACGTCGATTTCTGCTGAACTTGGAGGCATT	81
DB	1	CGATAACGATTGTGTGTGAGAGCGCAACGTCGATTTCTGCTGAACTTGGAGGCATT	60
QY	82	TCTACGACTTTCTCTCAGTCGAGGCTTTCTCCGACCCGATGCTCTTCAATTCGGTG	141
DB	61	TCTACGACTTTCTCTCAGTCGAGGCTTTCTCCGACCCGATGCTCTTCAATTCGGTG	120
QY	142	CTCCGCCAGCCAGCTTGGCGTCTTGAGAAATGGATGCTTTCACATACCCCTCTTCAA	201
DB	121	CTCCGCCAGCCAGCTTGGCGTCTTGAGAAATGGATGCTTTCACATACCCCTCTTCAA	180
QY	202	TCCCTTCTGACTGGTTATCAGTCAGTGGTAATGATGAACACACTTCTTATGGAGAAACA	261
DB	181	TCCCTTCTGACTGGTTATCAGTCAGTGGTAATGATGAACACACTTCTTATGGAGAAACA	240
QY	262	GGAGTCCGAGTTCTCTTTGGATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	321
DB	241	GGAGTCCGAGTTCTCTTTGGATGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	300
QY	322	CTAGCAGTTGCCAATGAAGAGCTTTGTTTCGATTTGTAACACAGAAATCAAAAGTTTC	381
DB	301	CTAGCAGTTGCCAATGAAGAGCTTTGTTTCGATTTGTAACACAGAAATCAAAAGTTTC	360
QY	382	AGAAAGAGTGTCTTAAAGATGGATGGCTCACTGGAATGCCGTCTTTGACCTGGCCTGG	441
DB	361	AGAAAGAGTGTCTTAAAGATGGATGGCTCACTGGAATGCCGTCTTTGACCTGGCCTGG	420
QY	442	GTTCTGTGTAACCTTAACTGTTACAGCAGAGGTGATCAACAGCCAAATTTGGGAC	501
DB	421	GTTCTGTGTAACCTTAACTGTTACAGCAGAGGTGATCAACAGCCAAATTTGGGAC	480
QY	502	GTAAAGCTGGTGAGCTGATTTGGAACATGCAAGGTTCATCAATGCAGCCTCAAGTCAGTT	561
DB	481	GTAAAGCTGGTGAGCTGATTTGGAACATGCAAGGTTCATCAATGCAGCCTCAAGTCAGTT	540
QY	562	GCCTTTTCTAAGTTTGAGAAAGCTGTATTCTGTACGGGTGGAAGAGATGGCAACATTATG	621
DB	541	GCCTTTTCTAAGTTTGAGAAAGCTGTATTCTGTACGGGTGGAAGAGATGGCAACATTATG	600
QY	622	GTCCTGGATACAGGTGCAACAAAAGATGGTTTATAGGCAAGTGAATCAATCAGT	681
DB	601	GTCCTGGATACAGGTGCAACAAAAGATGGTTTATAGGCAAGTGAATCAATCAGT	660
QY	682	GGAGCTCACATACTCAGACAGCAACCCCTTCAAAACCCCAAGAGAAACAGAAATCCA	741
DB	661	GGAGCTCACATACTCAGACAGCAACCCCTTCAAAACCCCAAGAGAAACAGAAATCCA	720
QY	742	AAAGGACTTGTCTCTGTGGATTTCCAGCAAGTGTACTGTGCTCTTTTCAAGAC	801
DB	721	AAAGGACTTGTCTCTGTGGATTTCCAGCAAGTGTACTGTGCTCTTTTCAAGAC	780
QY	802	GAGATACCTTAGTCTCAGCAGAGCTGTGGATGGGATTAATCAAGTATGGGATTTACGT	861
DB	781	GAGATACCTTAGTCTCAGCAGAGCTGTGGATGGGATTAATCAAGTATGGGATTTACGT	840
QY	862	AAGAAATTATCTGCTTATCGACAAGAACCCATAGCATCCAGTCTTCTCTGATCCAGGT	921

DB	841	AAGAATTATCTGCTTATCGACAAGAACCCATAGCATCCAAGTCTTTCTGCTACCCAGGT	900
QY	922	AGCAGACTCGAAACTTGGATATTCAAGTCGTGATTTGGATTCACCTGGCTCTACTTTA	981
DB	901	AGCAGACTCGAAACTTGGATATTCAAGTCGTGATTTGGATTCACCTGGCTCTACTTTA	960
QY	982	TTTGTCTAATTCACAGACGATAACATCTACATGTTTAAATATGACTGGGTGAAGACTTCT	1041
DB	961	TTTGTCTAATTCACAGACGATAACATCTACATGTTTAAATATGACTGGGTGAAGACTTCT	1020
QY	1042	CCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTATGTAAAAATCCAGCCTTAGT	1101
DB	1021	CCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTATGTAAAAATCCAGCCTTAGT	1080
QY	1102	CCAGATGACAGCTTTTATGTCAGTGGCTCAAGTGAAGCTGCCTACATATGGAAGGTC	1161
DB	1081	CCAGATGACAGCTTTTATGTCAGTGGCTCAAGTGAAGCTGCCTACATATGGAAGGTC	1140
QY	1162	TCACACCTCGCAACCTCTACTGTGCTCTCGGCTCATCTCAAGAGGTCAAGCTCTGTG	1221
DB	1141	TCACACCTCGCAACCTCTACTGTGCTCTCGGCTCATCTCAAGAGGTCAAGCTCTGTG	1200
QY	1222	TGCTGGTGTCCATCTGACTTCAAAAAGATGCTACTGTTCTGTATGACAATACTAAAA	1281
DB	1201	TGCTGGTGTCCATCTGACTTCAAAAAGATGCTACTGTTCTGTATGACAATACTAAAA	1260
QY	1282	ATCTGGCTTGAATAGAGGCTTAGAGGAGAAACAGAGGATGATAAACTTTCCAGGTC	1341
DB	1261	ATCTGGCTTGAATAGAGGCTTAGAGGAGAAACAGAGGATGATAAACTTTCCAGGTC	1320
QY	1342	GTTTGGGCTCTCAGAGAAAGAGCTCAAGACCTGGCTAGTAACAGTAAAGAGTAC	1401
DB	1321	GTTTGGGCTCTCAGAGAAAGAGCTCAAGACCTGGCTAGTAACAGTAAAGAGTAC	1380
QY	1402	CAGAGTACTCTCTGCCCCAAGCCCCCAGGGTAAAGTGAATCCATCCAAATTTTCCCCTCA	1461
DB	1381	CAGAGTACTCTCTGCCCCAAGCCCCCAGGGTAAAGTGAATCCATCCAAATTTTCCCCTCA	1440
QY	1462	TCCGAGCTTGTGCCCCAAGCTGTGCTGGAGACCTCCCTCTCTCTTCAAAATCTCTAGG	1521
DB	1441	TCCGAGCTTGTGCCCCAAGCTGTGCTGGAGACCTCCCTCTCTTCAAAATCTCTAGG	1500
QY	1522	TTTCTCTATTAACCTCTCTGCCCCAAGCCCCCAGGGTAAAGTGAATCCATCCAAATTTTCCCCTCA	1581
DB	1501	TTTCTCTATTAACCTCTCTGCCCCAAGCCCCCAGGGTAAAGTGAATCCATCCAAATTTTCCCCTCA	1560
QY	1582	TCCTCCGCTCTCTCCCAAGCCACTTTCATCTTTCAAGATGTGCGATTAGAAACTGGGTGACC	1641
DB	1561	TCCTCCGCTCTCTCCCAAGCCACTTTCATCTTTCAAGATGTGCGATTAGAAACTGGGTGACC	1620
QY	1642	CGAACCTCTCTCATACCCCAATCATCTCACTGCTTGGAGACCAAGATCATGTCT	1701
DB	1621	CGAACCTCTCTCATACCCCAATCATCTCACTGCTTGGAGACCAAGATCATGTCT	1680
QY	1702	CGGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCAATCCCAAGCAGAGGCTTGTCTGAG	1761
DB	1681	CGGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCAATCCCAAGCAGAGGCTTGTCTGAG	1740
QY	1762	TCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAG	1821
DB	1741	TCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAG	1800
QY	1822	TGTGTGAAGAGTTGAACTGTGTGACTGAGCTTGTATGGCAAGTTGAAATCTTCTATTG	1881
DB	1801	TGTGTGAAGAGTTGAACTGTGTGACTGAGCTTGTATGGCAAGTTGAAATCTTCTATTG	1860
QY	1882	GATCTGCTGCTGCTGGTAAACCCAGAGACCTTTAGTAAGGACTCTCTAGTCTCTACC	1941
DB	1861	GATCTGCTGCTGCTGGTAAACCCAGAGACCTTTAGTAAGGACTCTCTAGTCTCTACC	1920
QY	1942	MAATCAAGCAAAATTTGAAGGAGCTGGTATCTCAGAGCCTCCGCTCTCTATCAGT	2001

Qy	382	AGAAAGAGTGCTTCAAAGAATGGATGGCTCACTGGAAATCCGCTCTTTTGACCTGGCCTGG	441
Db	361	AGAAAGAGTGCTTCAAAGAATGGATGGCTCACTGGAAATCCGCTCTTTTGACCTGGCCTGG	420
Qy	442	GTTCCTCGTGAACCTTAAACTTTGTACAGCAGCAGTGATCAAAACAGACCAAAATTTTGGGAC	501
Db	421	GTTCCTCGTGAACCTTAAACTTTGTACAGCAGCAGTGATCAAAACAGACCAAAATTTTGGGAC	480
Qy	502	GTAAGAAGCTGGTGAGCTGATTTGGAAACATGCAAAAGTTCATCAATGCAGCCTCAAGTCAAGT	561
Db	481	GTAAGAAGCTGGTGAGCTGATTTGGAAACATGCAAAAGTTCATCAATGCAGCCTCAAGTCAAGT	540
Qy	562	GCCCTTTTCTAAGTTTGAAGAAAGCTGTATTCTGTACGGGTGGAAGAGATGGCAACATTATG	621
Db	541	GCCCTTTTCTAAGTTTGAAGAAAGCTGTATTCTGTACGGGTGGAAGAGATGGCAACATTATG	600
Qy	622	GTCTGGGATACACAGGTGCAACAAAAAAAGATGGGTTTTTATAGGCAAGTGAATCAAAATCAGT	681
Db	601	GTCTGGGATACACAGGTGCAACAAAAAAAGATGGGTTTTTATAGGCAAGTGAATCAAAATCAGT	660
Qy	682	GGAGCTCACAAATACCTCAGACAAGCAAAACCCCTTCAAACCCCAAGAAAGAACAGAAATTC	741
Db	661	GGAGCTCACAAATACCTCAGACAAGCAAAACCCCTTCAAACCCCAAGAAAGAACAGAAATTC	720
Qy	742	AAAGGACTTGCTCCTCTGTGGATTTCCAGCAAGTGTACTGTGTCTCTTTTCAAGAC	801
Db	721	AAAGGACTTGCTCCTCTGTGGATTTCCAGCAAGTGTACTGTGTCTCTTTTCAAGAC	780
Qy	802	GAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATTAATCAAAGTATGGGATTTACGT	861
Db	781	GAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATTAATCAAAGTATGGGATTTACGT	840
Qy	862	AAGAAATATATCTGCTTATCGACAAGAAACCCATAGCATCCAAAGTCTTTCCTGTACCCAGGT	921
Db	841	AAGAAATATATCTGCTTATCGACAAGAAACCCATAGCATCCAAAGTCTTTCCTGTACCCAGGT	900
Qy	922	AGCAGCACTCGAAACCTTGGATATTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTTA	981
Db	901	AGCAGCACTCGAAACCTTGGATATTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTTA	960
Qy	982	TTTGCTAATTGCAAGACGATAAACATCTACATGTTTAAATATGACTGGGTTGAAGACTTCT	1041
Db	961	TTTGCTAATTGCAAGACGATAAACATCTACATGTTTAAATATGACTGGGTTGAAGACTTCT	1020
Qy	1042	CCAGTGCTGCTATTTTCAATGGACACCGAACTCTACCTTTTATGTAATAATCCAGCCTTAGT	1101
Db	1021	CCAGTGCTGCTATTTTCAATGGACACCGAACTCTACCTTTTATGTAATAATCCAGCCTTAGT	1080
Qy	1102	CCAGATCAGCAGTTTTTTAGTTCAGTGGCTCAAGTGATGAAGCTGCCTACATATGGAAGTTC	1161
Db	1081	CCAGATCAGCAGTTTTTTAGTTCAGTGGCTCAAGTGATGAAGCTGCCTACATATGGAAGTTC	1140
Qy	1162	TCCACACCTTGGCAACCTCTACTGTGCTCTCTGGGTCAATTCCTCAAGAGGTCAAGTCTGTG	1221
Db	1141	TCCACACCTTGGCAACCTCTACTGTGCTCTCTGGGTCAATTCCTCAAGAGGTCAAGTCTGTG	1200
Qy	1222	TGCTGGTGTCCATCTGACTTCAAAAGATTGCTACCTGTTCTGTGATGACAATACACTAAAA	1281
Db	1201	TGCTGGTGTCCATCTGACTTCAAAAGATTGCTACCTGTTCTGTGATGACAATACACTAAAA	1260
Qy	1282	ATCTGGCGCTTGAATAGAGGCTTAGAGGAAACACGAGGTGTGATAAATTTTCCACGGTG	1341
Db	1261	ATCTGGCGCTTGAATAGAGGCTTAGAGGAAACACGAGGTGTGATAAATTTTCCACGGTG	1320
Qy	1342	GGTTGGGCTCTCAGAGAGAAAAAGAGTCAAGACCTGGCCTTAGTAAACAGTAGTACG	1401
Db	1321	GGTTGGGCTCTCAGAGAGAAAAAGAGTCAAGACCTGGCCTTAGTAAACAGTAGTACG	1380
Qy	1402	CAGAGTACTCTCTGCAAAAGCCCCCAGGGTAAAGTGCAATCCATCCAAATTTCTTCCCGTCA	1461
Db	1381	CAGAGTACTCTCTGCAAAAGCCCCCAGGGTAAAGTGCAATCCATCCAAATTTCTTCCCGTCA	1440

QY	1462	TCCGAGCTGTGTGCCCAAGCTGTGCTGAGAGACCTCCCTCTTCTCTTCAAAATACTCTCTACG	1521
DB	1441	TCCGAGCTGTGTGCCCAAGCTGTGCTGAGAGACCTCCCTCTTCTCTTCAAAATACTCTCTACG	1500
QY	1522	TTCTCTATTAATAAACCCTCTCTCGCCCAAGGCCCGGTCTCCCATCAACAGAGAGGCTCTGTCTC	1581
DB	1501	TTCTCTATTAATAAACCCTCTCTCGCCCAAGGCCCGGTCTCCCATCAACAGAGAGGCTCTGTCTC	1560
QY	1582	TCCTCCGCTCTCTCCCAAGCCACCTTCACTCTTTCAAGATGTGATTAAGAACTCGGTGACCC	1641
DB	1561	TCCTCCGCTCTCTCCCAAGCCACCTTCACTCTTTCAAGATGTGATTAAGAACTCGGTGACCC	1620
QY	1642	CGAACACTTCCTCATCACACCCATCACTCCACCTGCTTCGAGAGACCAAGATCATGTCTCT	1701
DB	1621	CGAACACTTCCTCATCACACCCATCACTCCACCTGCTTCGAGAGACCAAGATCATGTCTCT	1680
QY	1702	CCGAGAAAGCCCTTATTCTCTGTGAGCCAGAGTCATCCCAAGCAGAGGCTTGCTCTGTAG	1761
DB	1681	CCGAGAAAGCCCTTATTCTCTGTGAGCCAGAGTCATCCCAAGCAGAGGCTTGCTCTGTAG	1740
QY	1762	TCTAGAAATAGAGTAAGAGAGGAGCTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAAG	1821
DB	1741	TCTAGAAATAGAGTAAGAGAGGAGCTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAAG	1800
QY	1822	TGTGTGAAGAGTTGTAACTGTGTGACTGTGAGCTGTGATGAGCCAAAGTTGAAATCTTCATTG	1881
DB	1801	TGTGTGAAGAGTTGTAACTGTGTGACTGTGAGCTGTGATGAGCCAAAGTTGAAATCTTCATTG	1860
QY	1882	GATCTGTGCTTCCTGCTGTGCTTAACCAAGAGAGCTTAGTAAGAGACTCTCTAGTCTCTTACC	1941
DB	1861	GATCTGTGCTTCCTGCTGTGCTTAACCAAGAGAGCTTAGTAAGAGACTCTCTAGTCTCTTACC	1920
QY	1942	AAATCAAGCAAAATTTGAAGAGCTGGTATCACAGATATCTCAGAGCTCCGCTCTCTATCAGT	2001
DB	1921	AAATCAAGCAAAATTTGAAGAGCTGGTATCACAGATATCTCAGAGCTCCGCTCTCTATCAGT	1980
QY	2002	CCGTATGCTTCAGAAAGCTGTGAAACGCTACTCTCTCTTTTGAGACCTTGTGAGAGAGGG	2061
DB	1981	CCGTATGCTTCAGAAAGCTGTGAAACGCTACTCTCTCTTTTGAGACCTTGTGAGAGAGGG	2040
QY	2062	TCTGAAATGGTAGGCAAGAGATAGTTTCCCAAGAGNATAAAACTGGTTGTGGCCCATG	2121
DB	2041	TCTGAAATGGTAGGCAAGAGATAGTTTCCCAAGAGNATAAAACTGGTTGTGGCCCATG	2100
QY	2122	GCAGCAAAACGGAAGGCTGAGAAATCCATCTCCAAGAGTCCGTATCCCAAGACACCCAAT	2181
DB	2101	GCAGCAAAACGGAAGGCTGAGAAATCCATCTCCAAGAGTCCGTATCCCAAGACACCCAAT	2160
QY	2182	TCCAGGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTCAACATCAAGCCAGCTCCATG	2241
DB	2161	TCCAGGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTCAACATCAAGCCAGCTCCATG	2220
QY	2242	AGHAAATCTGCACATATCTTCATAGAAAGTCCAGGAGAGACTCTGTGTGCTCTGAACAC	2301
DB	2221	AGHAAATCTGCACATATCTTCATAGAAAGTCCAGGAGAGACTCTGTGTGCTCTGAACAC	2280
QY	2302	TCAACAGAAATATAGATTCTTAATCTGAGTCAGTTACTTGACTTTTGGTCCACTAAAAACAAG	2361
DB	2281	TCAACAGAAATATAGATTCTTAATCTGAGTCAGTTACTTGACTTTTGGTCCACTAAAAACAAG	2340
QY	2362	CTGAGCTTTGGTCCACTAAAAACAAGATGAAAAATACAAGAGTGACTCTATAACTCTGGTCT	2421
DB	2341	CTGAGCTTTGGTCCACTAAAAACAAGATGAAAAATACAAGAGTGACTCTATAACTCTGGTCT	2400
QY	2422	TTTAAGAAAGCTGCTTTTCAATTTTGTAGACAAATCTTTTCAACGCTGAAATGTACTTAA	2481
DB	2401	TTTAAGAAAGCTGCTTTTCAATTTTGTAGACAAATCTTTTCAACGCTGAAATGTACTTAA	2460
QY	2482	TCTGGTCTCTACTACCAATAAGTATGCGAGCTTCCCGAGGATCAATGCTGTGTTTAAAT	2541
DB	2461	TCTGGTCTCTACTACCAATAAGTATGCGAGCTTCCCGAGGATCAATGCTGTGTTTAAAT	2520
QY	2542	TCATAAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAAAATAT	2601

Db 2521 TCATAAGTAATTTGTCTACTAGCAATTTGAATGAATAGCTTCACTTTTAAATAT 2580
Qy 2602 TCATCTTCTCTATAATAATAGACATCCAGTTCATGGAGGCAAAAACAAGTTTCTTGTTA 2661
Db 2581 TCATCTTCTCTATAATAATAGACATCCAGTTCATGGAGGCAAAAACAAGTTTCTTGTTA 2640
Qy 2662 TCCTGAACTTTCTATGCTCAGTGGAAGTATCTGCCAGGCCACAGCATGAGGCCCTGTGAA 2721
Db 2641 TCCTGAACTTTCTATGCTCAGTGGAAGTATCTGCCAGGCCACAGCATGAGGCCCTGTGAA 2700
Qy 2722 GCGTGAAGTAAATCTCTGCTGAAGACCCCTGTTCTGTTCTGCTCCACATGTATA 2781
Db 2701 GCGTGAAGTAAATCTCTGCTGAAGACCCCTGTTCTGTTCTGCTCCACATGTATA 2760
Qy 2782 ATTTTATTTGAATAACATAATCTTTTCACTATG 2814
Db 2761 ATTTTATTTGAATAACATAATCTTTTCACTATG 2793
RESULT 14
ADO20231
ID ADO20231 standard; cDNA; 4221 BP.
XX ADO20231;
AC ADO20231;
XX
DT 12-AUG-2004 (first entry)
DE Human PRO polynucleotide #567.
XX
KW Human; PRO; gene; ss; immune related disorder;
KW systemic lupus erythematosus; rheumatoid arthritis; osteoarthritis;
KW juvenile chronic arthritis; systemic sclerosis; Sjogren's syndrome;
KW vasculitis; sarcoidosis; autoimmune haemolytic anaemia;
KW autoimmune thrombocytopenia; thyroiditis; diabetes mellitus;
KW renal disease; demyelinating disease; central nervous system;
KW peripheral nervous system; demyelinating polyneuropathy;
KW Guillain-Barre syndrome;
KW chronic inflammatory demyelinating polyneuropathy.
XX
OS Homo sapiens.
XX
XX WO2004043361-A2.
XX
XX 27-MAY-2004.
XX
XX 06-NOV-2003; 2003WO-US035268.
XX
XX 08-NOV-2002; 2002US-0425235P.
XX
XX (GETH) GENENTECH INC.
XX
XX Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
XX Wood WJ, Wu TD;
XX
XX WPI; 2004-420067/39.
XX
XX P-PSDB; ADO20232.
XX
XX Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
XX treating an immune related disorder such as systemic lupus erythematosus,
XX rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
XX spondyloarthropathy.
XX
XX Claim 1; SEQ ID NO 1202; 1731pp; English.
XX
XX The invention relates to human PRO polypeptides and the polynucleotides
XX encoding them. The polypeptides and polynucleotides are useful for
XX treating and diagnosing immune related disorders in mammals. The immune
XX related disorders include systemic lupus erythematosus, rheumatoid
XX arthritis, osteoarthritis, juvenile chronic arthritis, systemic
XX sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
XX haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
XX mellitus, immune-mediated renal disease, demyelinating diseases of the

CC central or peripheral nervous system, demyelinating polyneuropathy,
CC Guillain-Barre syndrome and chronic inflammatory demyelinating
CC polyneuropathy. This sequence represents a human PRO polynucleotide of
CC the invention.
XX
SQ Sequence 4221 BP; 1229 A; 897 C; 846 G; 1249 T; 0 U; 0 Other;
Query Match 98.5%; Score 2789.8; DB 12; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 22 CGATAACGATTGTGTGAGAGCGCAACGTCGCGATTTCTGCTGAACATTTGGAGGCATT 81
Db 1 CGATAACGATTGTGTGAGAGCGCAACGTCGCGATTTCTGCTGAACATTTGGAGGCATT 60
Qy 82 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCGACCCCTGATGCTTCAATTCGGTG 141
Db 61 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCGACCCCTGATGCTTCAATTCGGTG 120
Qy 142 CTCGCCAGCCCGAGCTTGGCGTCTGAGAAATGAGTGGTCTTCAATACCTCTTCAA 201
Db 121 CTCGCCAGCCCGAGCTTGGCGTCTGAGAAATGAGTGGTCTTCAATACCTCTTCAA 180
Qy 202 TCCCTTCTGACTGTTATCAGTCAGTGGTAAATGATGAACACACTTCTTATGAGAAACA 261
Db 181 TCCCTTCTGACTGTTATCAGTCAGTGGTAAATGATGAACACACTTCTTATGAGAAACA 240
Qy 262 GGAGTCCGAGTTCCTCTTTTGGATGTAACCTTCTCTCTGCTCCCAATATGGAACATGTA 321
Db 241 GGAGTCCGAGTTCCTCTTTTGGATGTAACCTTCTCTCTGCTCCCAATATGGAACATGTA 300
Qy 322 CTAGCAGTTGCCAATGAAGAAGGCTTTGTCGATGTATACACAGATCAACAAGTTTC 381
Db 301 CTAGCAGTTGCCAATGAAGAAGGCTTTGTCGATGTATACACAGATCAACAAGTTTC 360
Qy 382 AGAAGAAGTCTCTCAAGAATGATGGCTCACTGGAATGCCGTCTTGGACTGCGCTGG 441
Db 361 AGAAGAAGTCTCTCAAGAATGATGGCTCACTGGAATGCCGTCTTGGACTGCGCTGG 420
Qy 442 GTTCTGCTGAACTTAAACTTTGTTACAGCAGCAGGTGATCAAAACAGCCAAATTTGGAC 501
Db 421 GTTCTGCTGAACTTAAACTTTGTTACAGCAGCAGGTGATCAAAACAGCCAAATTTGGAC 480
Qy 502 GTAAAGCTGTGAGCTGATTGGAAACATGCAAGGTCAATCAATGCAGCTCAAGTCAGTT 561
Db 481 GTAAAGCTGTGAGCTGATTGGAAACATGCAAGGTCAATCAATGCAGCTCAAGTCAGTT 540
Qy 562 GCCTTTTCTAAGTTTGAAAAAGCTGTATTCTGTACGGGTGGAAGAGATGGCAACATTATG 621
Db 541 GCCTTTTCTAAGTTTGAAAAAGCTGTATTCTGTACGGGTGGAAGAGATGGCAACATTATG 600
Qy 622 GTCTGGGATACAGGTGCAACAAAAAGATGGTTTTATAGGCAAGTGAATCAAAATCAGT 681
Db 601 GTCTGGGATACAGGTGCAACAAAAAGATGGTTTTATAGGCAAGTGAATCAAAATCAGT 660
Qy 682 GGAGCTCAAAATACCTCAGACAGCAAAACCCCTTCAAAACCCCAAGAAACAGAAATTC 741
Db 661 GGAGCTCAAAATACCTCAGACAGCAAAACCCCTTCAAAACCCCAAGAAACAGAAATTC 720
Qy 742 AAAGGACTTGTCTCTTCTGTGGATTTCCAGCAAAAGTGTACTGTGGTCTCTTTCAAGAC 801
Db 721 AAAGGACTTGTCTCTTCTGTGGATTTCCAGCAAAAGTGTACTGTGGTCTCTTTCAAGAC 780
Qy 802 GAGATACCTTAGTCTCAGCAGGAGCTGTGGATGGATAAATCAAAAGTATGGGATTTAGT 861
Db 781 GAGATACCTTAGTCTCAGCAGGAGCTGTGGATGGATAAATCAAAAGTATGGGATTTAGT 840
Qy 862 AAGATTTACTGCTTATTCGACNAGAACCCATAGATCCAAAGTCTTCTCTGACCCAGT 921
Db 841 AAGATTTACTGCTTATTCGACNAGAACCCATAGATCCAAAGTCTTCTCTGACCCAGT 900
Qy 922 AGCAGCACTCGAAAACTTGGATATTTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 981

901 AGCAGCACTCGAAAACTTGGATATTCAAGTCTGATTTTGGATTCCACTGGCTCTACTTTTA 960
982 TTTTGCTAAATTGCACAGCACTACACTACATGTTTAAATATGACTGGGTGAGACTTCT 1041
961 TTTTGCTAAATTGCACAGCACTACACTACATGTTTAAATATGACTGGGTGAGACTTCT 1020
1042 CCAGTGGCTATTTTCAATGGACACCAAGAACTCTACCTTTTATGTAAAAATCCAGCCCTTAGT 1101
1021 CCAGTGGCTATTTTCAATGGACACCAAGAACTCTACCTTTTATGTAAAAATCCAGCCCTTAGT 1080
1102 CCAGATGACCAAGTTTTAGTCACTGGCTCAAGTGATGAAGCTGCTACATATGGAAGGTC 1161
1081 CCAGATGACCAAGTTTTAGTCACTGGCTCAAGTGATGAAGCTGCTACATATGGAAGGTC 1140
1162 TCCACACCTGGCAACCTCTACTGCTGCTCTGGGTCAATCTCAAGAGGTCAGCTCTGTG 1221
1141 TCCACACCTGGCAACCTCTACTGCTGCTCTGGGTCAATCTCAAGAGGTCAGCTCTGTG 1200
1222 TGTGCTGTCTCATCTGACTTTCACAAAGATTGCTTACCTGTTCTGATGACAAATACACTAAAA 1281
1201 TGTGCTGTCTCATCTGACTTTCACAAAGATTGCTTACCTGTTCTGATGACAAATACACTAAAA 1260
1282 ATCTGCGCTTGAATAGAGCTTTAGAGGAGAAACCAAGAGGTTGATAAAACTTTCCACGGTG 1341
1261 ATCTGCGCTTGAATAGAGCTTTAGAGGAGAAACCAAGAGGTTGATAAAACTTTCCACGGTG 1320
1342 GGTGGGCTCTCAGAAAGAAAAGATCAAGACCTGGCTAGTACAGTAACGAGTAGC 1401
1321 GGTGGGCTCTCAGAAAGAAAAGATCAAGACCTGGCTAGTAAACAGTAACGAGTAGC 1380
1402 CAGAGTACTCTCCGCAAGACCCCAAGGTAAGTGCAATCCATCCAAATCTTCCCGGTCA 1461
1381 CAGAGTACTCTCCGCAAGACCCCAAGGTAAGTGCAATCCATCCAAATCTTCCCGGTCA 1440
1462 TCCGAGCTTGTGCCCCAAGCTGTGCTGGAGACCTCCCTCTTCCCTTCAAAATACCTCTACG 1521
1441 TCCGAGCTTGTGCCCCAAGCTGTGCTGGAGACCTCCCTCTTCCCTTCAAAATACCTCTACG 1500
1522 TTCTCTATTAAACCTCTCTGCAAGCCCACTTCATCTTCAAGATGCGAATAGAACTGGGTGACC 1641
1561 TCTCTCGTCTCTCCCAAGCACCTTCATCTTCAAGATGCGAATAGAACTGGGTGACC 1620
1642 CGAACACCTTCTCTCATCACCAACCCATCACTCCACCTGCTTCCGAGACCAAGATCATGTCT 1701
1621 CGAACACCTTCTCTCATCACCAACCCATCACTCCACCTGCTTCCGAGACCAAGATCATGTCT 1680
1702 CGAGAAAAAGCCCTTATTTCTGTGAGCGAGAGTCAATCCCAAGCAGAGGCTTGTCTGAG 1761
1681 CGAGAAAAAGCCCTTATTTCTGTGAGCGAGAGTCAATCCCAAGCAGAGGCTTGTCTGAG 1740
1762 TCTAGAAAAATAGCTAAAGAGGCGCTAGACTCAAGCTGTCTGGAGAGTGTGAAAAAAG 1821
1741 TCTAGAAAAATAGCTAAAGAGGCGCTAGACTCAAGCTGTCTGGAGAGTGTGAAAAAAG 1800
1822 TGTGTGAAGAGTTGTAACCTGTGTGACTGAGCTTGTATGGCCAAAGTTGAAAAATCTTCAATTTG 1881
1801 TGTGTGAAGAGTTGTAACCTGTGTGACTGAGCTTGTATGGCCAAAGTTGAAAAATCTTCAATTTG 1860
1882 GATCTGTGCTGCTTGTGCTGTAAACCAAGAGACCTTAGTAAGACTCTCTAGGTCTCTACC 1941
1861 GATCTGTGCTGCTTGTGCTGTAAACCAAGAGACCTTAGTAAGACTCTCTAGGTCTCTACC 1920
1942 AAATCAAGCAAAATTCAGAGGCTGTACCAGTATCTCAGAGCCTCCGCTCTCTATCAGT 2001
1921 AAATCAAGCAAAATTCAGAGGCTGTGTACCAAGTATCTCAGAGCCTCCGCTCTCTATCAGT 1980
2002 CCGTATGCTTTCAGAAAGCTGTGGAACGCTACCTCTTCCCTTTGAGACCTTGTGGAGAGGG 2061
1981 CCGTATGCTTTCAGAAAGCTGTGGAACGCTACCTCTTCCCTTTGAGACCTTGTGGAGAGGG 2040

2062 TCTGAAATGGTAGGCAAGAGAAATAGTTCCCCAGAGAAATAAAACTGGTGTGTTGGCCATG 2121
2041 TCTGAAATGGTAGGCAAGAGAAATAGTTCCCCAGAGAAATAAAACTGGTGTGTTGGCCATG 2100
2122 GCAGCCAAACGGAAGGCTGAGAATCCATCTCCACGAAGTCCGTCAATCCAGACACCCCAAT 2181
2101 GCAGCCAAACGGAAGGCTGAGAATCCATCTCCACGAAGTCCGTCAATCCAGACACCCCAAT 2160
2182 TCCAGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTCCACCATCAGCCCGAGCTCCATG 2241
2161 TCCAGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTCCACCATCAGCCCGAGCTCCATG 2220
2242 AGAAAAATCTGCACATACTTCCATAGAAAGTCCCAAGAGGACTTCTGTGGTCTCTGAACAC 2301
2221 AGAAAAATCTGCACATACTTCCATAGAAAGTCCCAAGAGGACTTCTGTGGTCTCTGAACAC 2280
2302 TCAACAGAAATATAGATTTAAATCTGAGTGAGTACTAGAGCTTTGGTCCACTAAACCAAG 2361
2281 TCAACAGAAATATAGATTTAAATCTGAGTGAGTACTAGAGCTTTGGTCCACTAAACCAAG 2340
2362 CTGAGCTTTGGTCCACTAAACAGATGAAAAATACAGAGTACTCTATAACTCTGGTC 2421
2341 CTGAGCTTTGGTCCACTAAACAGATGAAAAATACAGAGTACTCTATAACTCTGGTC 2400
2422 TTTTAAAGAAAGCTGCCCTTTTTCATTTTATAGACAAAAATCTTTTCAACGCTGAAATGTACTAA 2481
2401 TTTTAAAGAAAGCTGCCCTTTTTCATTTTATAGACAAAAATCTTTTCAACGCTGAAATGTACTAA 2460
2482 TCTGTGTTCTACTACCATATATGATGACAGCTTCCCGAGGATGAATGCTGTGTTAAATTT 2541
2461 TCTGTGTTCTACTACCATATATGATGACAGCTTCCCGAGGATGAATGCTGTGTTAAATTT 2520
2542 TCATAAGGTAATTTGTGCATCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTTAT 2601
2521 TCATAAGGTAATTTGTGCATCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTTAT 2580
2602 TCATCTCTCTATAATAATGACATCCAGTTCATCGAGGCAAAAAACAAGTTTCTTGTTA 2661
2581 TCATCTCTCTATAATAATGACATCCAGTTCATCGAGGCAAAAAACAAGTTTCTTGTTA 2640
2662 TCTGAAACCTTTCTATGCTCAGTGGAAAGTATCTGCCAGCCACAGCATGAGGCTGTGAA 2721
2641 TCTGAAACCTTTCTATGCTCAGTGGAAAGTATCTGCCAGCCACAGCATGAGGCTGTGAA 2700
2722 GGTGACTGAGAAATCTCTGCTGAAAGCCCTGGTCTGTTCTGCTTCTGCTCCAAACATGTATA 2781
2701 GGTGACTGAGAAATCTCTGCTGAAAGCCCTGGTCTGTTCTGCTTCTGCTCCAAACATGTATA 2760
2782 ATTTTATTTGAAATACATAATCTTTTCACTATG 2814
2761 ATTTTATTTGAAATACATAATCTTTTCACTATG 2793

RESULT 15

ADR25784
ID ADR25784 standard; DNA; 4221 BP.

XX ADR25784;

DT 21-OCT-2004 (first entry)

DE Breast cancer prognosis marker #1645.

ds; breast cancer; prognosis; gene expression; diagnosis.

OS Homo sapiens.

XX WO2004065545-A2.

XX 05-AUG-2004.

XX 15-JAN-2004; 2004WO-US001100.

PF

XX 15-JAN-2003; 2003US-00342887.
XX (ROSE-) ROSETTA INPHARMATICS LLC.
PA (NECA-) NETHERLANDS CANCER INST.
XX
PI Van't Veer LJ, He Y;
DR WPI; 2004-593473/57.
XX
XX Classifying a breast cancer patient according to prognosis comprises
PT determining the similarity between the level of expression of each of
PT five genes in a cell sample taken from patient, to control levels.
XX
PS Disclosure; SEQ ID NO 1645; 226pp; English.
XX
XX The invention relates to a method of classifying a breast cancer patient
CC according to prognosis by determining the similarity between the level of
CC expression of each of five genes for which markers are listed in the
CC specification, in a cell sample taken from the breast cancer patient, to
CC control levels of expression for each respective five genes to obtain a
CC patient similarity value. The methods are useful for classifying a breast
CC cancer patient according to prognosis. Kits and computer program products
CC are useful for data analysis using the diagnostic, prognostic and
CC statistical methods of the invention. This sequence corresponds to a
CC marker used in the method of the invention.
XX
XX Sequence 4221 BP; 1229 A; 897 C; 846 G; 1249 T; 0 U; 0 Other;
Query Match 98.5%; Score 2789.8; DB 13; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 22 CGATAACGATTGTTGTGTCAGAGGCGCAACGTGCGATTTCCTGCTGAACTTGGAGGCATT 81
DB 1 CGATAACGATTGTTGTGTCAGAGGCGCAACGTGCGATTTCCTGCTGAACTTGGAGGCATT 60
QY 82 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTGCGGTG 141
DB 61 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTGCGGTG 120
QY 142 CTCGCCGACCCGCTTGGCGTCTGAGAAATGGATGGTCTTCAATACCCCTCTTCAA 201
DB 121 CTCGCCGACCCGCTTGGCGTCTGAGAAATGGATGGTCTTCAATACCCCTCTTCAA 180
QY 202 TCCCTTCTGACTGGTTATCAGTGCAGTGTGTAATGATGAACACACTTCTTATGGAGAAACA 261
DB 181 TCCCTTCTGACTGGTTATCAGTGCAGTGTGTAATGATGAACACACTTCTTATGGAGAAACA 240
QY 262 GGAGTCCCAAGTTCCTCTTTTGGATGATCCTTCTCTTCTGCTCCCAATATGAAACATGTA 321
DB 241 GGAGTCCCAAGTTCCTCTTTTGGATGATCCTTCTCTTCTGCTCCCAATATGAAACATGTA 300
QY 322 CTAGCAGTTGCCAATGAAGAGCTTTGTTGCGATTGTATACACAGAAATCAAGAGTTTC 381
DB 301 CTAGCAGTTGCCAATGAAGAGCTTTGTTGCGATTGTATACACAGAAATCAAGAGTTTC 360
QY 382 AGAAGAGTGTCTCAAGAAATGGATGGCTCACTGGAATGCCGTCTTTGACCTGGCCTGG 441
DB 361 AGAAGAGTGTCTCAAGAAATGGATGGCTCACTGGAATGCCGTCTTTGACCTGGCCTGG 420
QY 442 GTTCTCGTGAACCTTAAACTTTGTTACAGCAGAGGTGATCAACAGCCAAAATTTGGGAC 501
DB 421 GTTCTCGTGAACCTTAAACTTTGTTACAGCAGAGGTGATCAACAGCCAAAATTTGGGAC 480
QY 502 GTAAAGCTGGTGCAGCTGATGGAAATCAACGAAGGTTCATCAATGACGCTCAAGTCAGTT 561
DB 481 GTAAAGCTGGTGCAGCTGATGGAAATCAACGAAGGTTCATCAATGACGCTCAAGTCAGTT 540
QY 562 GCCTTTTCTAAAGTTTGAGAAAGCTGTATTCTGTACGGGTGGAGAGATGGCAACATTATG 621
DB 541 GCCTTTTCTAAAGTTTGAGAAAGCTGTATTCTGTACGGGTGGAGAGATGGCAACATTATG 600

QY 622 GTCTGGGATACAGGTGCAACAAAAGATGCGTTTTTATAGCAAGTGAATCAAAATCAGT 681
DB 601 GTCTGGGATACAGGTGCAACAAAAGATGCGTTTTTATAGCAAGTGAATCAAAATCAGT 660
QY 682 GGAGCTCACAATACCTCAGACAAGCAAAACCCCTTCAAAACCCCAAGAAACAGAAATTC 741
DB 661 GGAGCTCACAATACCTCAGACAAGCAAAACCCCTTCAAAACCCCAAGAAACAGAAATTC 720
QY 742 AAAGGACTTGTCTCTCTGTGGATTTCAGAAAAGTGTACTGTGGTCTCTTTCAAGAC 801
DB 721 AAAGGACTTGTCTCTCTGTGGATTTCAGAAAAGTGTACTGTGGTCTCTTTCAAGAC 780
QY 802 GAGATACCTTGTAGTCTCAGCAGGAGCTGTGATGCGATAAATCAAAAGTATGGGATTTAGT 861
DB 781 GAGATACCTTGTAGTCTCAGCAGGAGCTGTGATGCGATAAATCAAAAGTATGGGATTTAGT 840
QY 862 AAGAAATTAATCTGCTTATTCGACAAGAACCCATAGCATCCAAGTCTTTCCTGTACCAGT 921
DB 841 AAGAAATTAATCTGCTTATTCGACAAGAACCCATAGCATCCAAGTCTTTCCTGTACCAGT 900
QY 922 AGCAGACTCGAAAACCTTGGATATTCAGAGTCTGATTTTGGATTCACATGGCTCTACTTTA 981
DB 901 AGCAGACTCGAAAACCTTGGATATTCAGAGTCTGATTTTGGATTCACATGGCTCTACTTTA 960
QY 982 TTTGCTAATTCGACAGACGATAACATCTACATGTTTAAATATGACTGGGTGGAAGACTTCT 1041
DB 961 TTTGCTAATTCGACAGACGATAACATCTACATGTTTAAATATGACTGGGTGGAAGACTTCT 1020
QY 1042 CCAGTGGCTATTTTCAATGGACACACAGAACTCTACTTTTATGTAATAATCCAGCCTTAGT 1101
DB 1021 CCAGTGGCTATTTTCAATGGACACACAGAACTCTACTTTTATGTAATAATCCAGCCTTAGT 1080
QY 1102 CCAGATGACAGTTTTATAGTCAAGTGGCTCAAGATGATGAAGTGCCTACATATGGAAGTTC 1161
DB 1081 CCAGATGACAGTTTTATAGTCAAGTGGCTCAAGATGATGAAGTGCCTACATATGGAAGTTC 1140
QY 1162 TCCACACCTCGCACCTCTACTGTGCTCTCGGTGCTTCTCAAGAGTCAAGTCTGTG 1221
DB 1141 TCCACACCTCGCACCTCTACTGTGCTCTCGGTGCTTCTCAAGAGTCAAGTCTGTG 1200
QY 1222 TGCTGGTCTCATCTGACTTTCACAAAGATTGCTACTCTTCTGATGACAATACATAA 1281
DB 1201 TGCTGGTCTCATCTGACTTTCACAAAGATTGCTACTCTTCTGATGACAATACATAA 1260
QY 1282 ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAAACAGAGAGTGAATAAATTTCCACGGTG 1341
DB 1261 ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAAACAGAGAGTGAATAAATTTCCACGGTG 1320
QY 1342 GGTGGGCTCTCAGAAAGAAAAGAGTCAAGACCTGGCCTAGTAAACAGTAAACAGTACG 1401
DB 1321 GGTGGGCTCTCAGAAAGAAAAGAGTCAAGACCTGGCCTAGTAAACAGTAAACAGTACG 1380
QY 1402 CAGAGTACTCTCTGCAAGGCTTAAAGTGAATTCATCCATCCAAATTTTCCCGCTCA 1461
DB 1381 CAGAGTACTCTCTGCAAGGCTTAAAGTGAATTCATCCATCCAAATTTTCCCGCTCA 1440
QY 1462 TCCGAGCTTGTGCCCCAAGCTGTGCTGGAGACCTCTCTTCTTCAAAATCTCTAG 1521
DB 1441 TCCGAGCTTGTGCCCCAAGCTGTGCTGGAGACCTCTCTTCTTCAAAATCTCTAG 1500
QY 1522 TTCTCTATTAACCTCTCTGCAAGGCTTCCCATCAACAGAGAGGCTCTGTG 1581
DB 1501 TTCTCTATTAACCTCTCTGCAAGGCTTCCCATCAACAGAGAGGCTCTGTG 1560
QY 1582 TCCTCCGTCTCTCCCAAGCCACTTTCATCTTTCAAGATGTGCGATTAGAAATCTGGGTGACC 1641
DB 1561 TCCTCCGTCTCTCCCAAGCCACTTTCATCTTTCAAGATGTGCGATTAGAAATCTGGGTGACC 1620
QY 1642 CGAAACCTTCTCATCCACCCATCATCTCCACTGCTTGGAGACCAAGATCATGTCT 1701
DB 1621 CGAAACCTTCTCATCCACCCATCATCTCCACTGCTTGGAGACCAAGATCATGTCT 1680
QY 1702 CCGAGAAAAGCCCTTATTTCTCTGTGAGCCAGAGTCAATCCCAAGCAGAGGCTTGTCTGTAG 1761

```
Db 1681 ||||| CCAGAAAAAGCCCTTATCTCTGTGAGCCAGAGTCATCCCAAGCAGAGGCTTCTCTGAG 1740
Qy 1762 TCTAGAAATAGATAAGAGGAGCTAGACTCAAGCTGTCTGAGAGTGTGAACAAG 1821
Db 1741 TCTAGAAATAGATAAGAGGAGCTAGACTCAAGCTGTCTGAGAGTGTGAACAAG 1800
Qy 1822 TGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGTATGGCCAAAGTGTGAATACTTCAATTTG 1881
Db 1801 TGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGTATGGCCAAAGTGTGAATACTTCAATTTG 1860
Qy 1882 GATCTGTGCTGCTTGTCTGTGTAAACAGGAAGACTTCTAGTAAGAGCTCTCTAGTCTCTACC 1941
Db 1861 GATCTGTGCTGCTTGTCTGTGTAAACAGGAAGACTTCTAGTAAGAGCTCTCTAGTCTCTACC 1920
Qy 1942 AATCAAGCAAAATTCAGAGAGCTGTGTACAGTATCTCAGAGCTCGGTCTCCTATCAGT 2001
Db 1921 AATCAAGCAAAATTCAGAGAGCTGTGTACAGTATCTCAGAGCTCGGTCTCCTATCAGT 1980
Qy 2002 CCGTATGCTTCAGAAAGCTGTGAAACGCTACCTCTCTCTTTGAGACCTTGTGGAGAAGG 2061
Db 1981 CCGTATGCTTCAGAAAGCTGTGAAACGCTACCTCTCTCTTTGAGACCTTGTGGAGAAGG 2040
Qy 2062 TCTGAAATGCTAGGCAAGAGAAATAGTTCCTCCAGAGAAATAAAACTGGTTGTTGGCCATG 2121
Db 2041 TCTGAAATGCTAGGCAAGAGAAATAGTTCCTCCAGAGAAATAAAACTGGTTGTTGGCCATG 2100
Qy 2122 GCAGCCAAACGGAAGCTGAGAAATCCATCTCCAGGAAGTCCGTCTCCAGAGACCCCAAT 2181
Db 2101 GCAGCCAAACGGAAGCTGAGAAATCCATCTCCAGGAAGTCCGTCTCCAGAGACCCCAAT 2160
Qy 2182 TCCAGGAGACAGAGCGGAAGACATTCGCAAGCCGCTCACCATCAGCCAGCTCCATG 2241
Db 2161 TCCAGGAGACAGAGCGGAAGACATTCGCAAGCCGCTCACCATCAGCCAGCTCCATG 2220
Qy 2242 AGGAAATCTGCACATACATTTCCATAGAAAGTCCAGGAGGACTTCTGTGTCTCTGAACAC 2301
Db 2221 AGGAAATCTGCACATACATTTCCATAGAAAGTCCAGGAGGACTTCTGTGTCTCTGAACAC 2280
Qy 2302 TCACAGAAATATAGATTTCTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTAAACAAAG 2361
Db 2281 TCACAGAAATATAGATTTCTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTAAACAAAG 2340
Qy 2362 CTGAGCTTTGGTCCACTAAACAAAGATGAAATACAAAGTGAGCTCTATACTCTGGTC 2421
Db 2341 CTGAGCTTTGGTCCACTAAACAAAGATGAAATACAAAGTGAGCTCTATACTCTGGTC 2400
Qy 2422 TTTAAGAAAGCTGCCCTTTTCATTTTGTAGACAAATCTTTTCAACGCTGAAATGTACTAA 2481
Db 2401 TTTAAGAAAGCTGCCCTTTTCATTTTGTAGACAAATCTTTTCAACGCTGAAATGTACTAA 2460
Qy 2482 TCTGGTTCTATACCAATATATATATGACGTTCCCGAGATGAATGCTGTGTTAAAT 2541
Db 2461 TCTGGTTCTATACCAATATATATATGACGTTCCCGAGATGAATGCTGTGTTAAAT 2520
Qy 2542 TCATAAGTAATTTCTCTAGCTCTAGCATTTTGAATGAAATAGTCTTCACTTTTAAATTA 2601
Db 2521 TCATAAGTAATTTCTCTAGCTCTAGCATTTTGAATGAAATAGTCTTCACTTTTAAATTA 2580
Qy 2602 TCATCTTCTATATAATATGACATCCAGTTTCATGGAGGCAAAAAACAAGTTCTTGTTA 2661
Db 2581 TCATCTTCTATATAATATGACATCCAGTTTCATGGAGGCAAAAAACAAGTTCTTGTTA 2640
Qy 2662 TCCGAAACTTCTATGCTCAGTGAAGATATCTGCCAGCCAGCATGAGGCTGTGAA 2721
Db 2641 TCCGAAACTTCTATGCTCAGTGAAGATATCTGCCAGCCAGCATGAGGCTGTGAA 2700
Qy 2722 GGCTGACTGAGAAATCCTCTGCTGAGAGCCCTGGTTCTGTTCTGCTCCACATGTATA 2781
Db 2701 GGCTGACTGAGAAATCCTCTGCTGAGAGCCCTGGTTCTGTTCTGCTCCACATGTATA 2760
Qy 2782 ATTTTATTTGAAATACATAATCTTTTCACTATG 2814
|||||
```

Db 2761 ATTTTATTTGAAATACATAATCTTTTCACTATG 2793

Search completed: November 9, 2006, 06:21:28
Job time : 1693 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2006, 18:08:54 ; Search time 46 Seconds
(without alignments)
1526.918 Million cell updates/sec

Title: US-10-726-160-2

Perfect score: 3877

Sequence: 1 MLFNSVLRQPQLGVLNGWS.....TYFHRKSQDFCGPEHSTEL 730

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 80.*

1: Pirl.*

2: Pirl.*

3: Pirl.*

4: Pirl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	793.5	20.5	758	S51748	lethal(2)denticlel
2	370.5	9.6	727	T24284	hypothetical prote
3	340	8.0	810	T48835	lethal(2)denticlel
4	295.5	7.6	1526	AC2239	WD-40 repeat prote
5	286.5	7.4	490	T37884	transcription fact
6	279.5	7.2	934	AG1989	WD-40 repeat prote
7	271	7.0	676	AH2195	hypothetical prote
8	268	6.9	1189	AH2154	WD-repeat protei
9	261.5	6.7	1227	AC1810	WD-40 repeat prote
10	260	6.7	1747	AC1842	WD-40 repeat prote
11	259.5	6.7	1356	T18521	beta transducin-li
12	258.5	6.7	1258	AI2155	WD-repeat protei
13	258.5	6.7	1708	AI2186	WD-40 repeat prote
14	256.5	6.6	343	C84870	probable splicing
15	256	6.6	515	S19487	hypothetical prote
16	254.5	6.6	1683	AF2071	WD-40 repeat prote
17	252	6.5	1189	AI2493	WD-repeat protei
18	250	6.4	1711	AD1842	WD-40 repeat prote
19	248.5	6.4	1551	AB2410	WD-repeat protei
20	247.5	6.4	303	S45461	hypothetical prote
21	236.5	6.1	559	AB2202	hypothetical prote
22	235.5	6.1	1049	T42045	beta transducin-li
23	234	6.0	317	T46032	WD-40 repeat regul
24	233.5	6.0	606	T11352	WDrl protein - hum
25	232.5	6.0	342	AE2490	WD-repeat protei
26	230.5	5.9	677	AE1861	serine/threonine k
27	227.5	5.9	340	T39362	u5 snRNP-like prot
28	225.5	5.8	473	T33805	hypothetical prote
29	225.5	5.8	786	AG2375	WD-40 repeat-prote

RESULT 1

S51748

lethal(2)denticleless - fruit fly (Drosophila melanogaster)

C;Species: Drosophila melanogaster

C;Date: 07-May-1995 #sequence_revision 01-Sep-1995 #text_change 05-Oct-2004

C;Accession: S51748

R;Kurzik-Dumke, U.; Neubauer, M.; Debes, A.

submitted to the EMBL Data Library, December 1994

A;Description: Identification of the novel Drosophila melanogaster heat shock gene, leth

A;Reference number: S51748

A;Accession: S51748

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-758 <KUR>

A;Cross-references: UNIPROT:Q24371; UNIPARC:UPI000012E18D; EMBL:X83414; NID:G603538; PID

C;Genetics:

A;Gene: FlyBase:1(2)dt1

A;Cross-references: FlyBase:FBgn0013548

A;Introns: 3/2

F;206-239/Domain: WD repeat homology <WD1>

F;349-383/Domain: WD repeat homology <WD2>

Query Match 20.5%; Score 793.5; DB 2; Length 758;

Best Local Similarity 29.2%; Pred. No. 1.7e+40;

Matches 224; Conservative 112; Mismatches 268; Indels 163; Gaps 25;

QY 49 PVPP-FGCTFSSAPNMEHVLAVANEFGFVRLYNT-----ESQSF-RKKCFKEMAHWN 99

Db 38 PEPPISAKFANCDGYRHILAIANEDGKITLQDTTQRNHQPERQSLVGPQC-----HPN 91

QY 100 AVFDLAWFGELKLVTAAGDQTAKFWDVWAGELIG--TKGHQCSLKSVAFSFEKAVFC 157

Db 92 AVFDLEWAPQWRFFVSASGDHTARLMEVAGSGIRGLNSYVGHTRSVKSAFRTDPAVFA 151

QY 158 TGRDGNIMVWTRCNKDDGFYQVQNIQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFFQ 217

Db 152 TGRDGAIIWDIRANLNMDLTSRVDCNYSGHTGGPGPIPVSKORKTRTPKAWAGTTSS 211

QY 218 SVTVVLFDENTILVSAVADGIIKVDLRLKNTAYROEPIASKSFLYPGSSSTRKLYSSL 277

Db 212 SITGLAFQNDTLISCGAGDGVIKWDLARNYATYKKEPLRPHKLPYAGSSTFR-GFTNL 270

QY 278 ILDTSGSTLIFANCTDDNIYFNWMTGLKSPVAIFNGHQNSTFFVKVSSLSGDDFLVSGSS 337

Db 271 IVNASGTRLYANGMDNTIYCNLASYSRPLACYKGLNLSFTFYIKSCLSPDGKYLSSGS 330

QY 338 DEAYITWYKSTPWPQPTVLLGHSEVTSVCMCPSEDTTKIATCDDNTLKIWLNL----- 391

Db 331 DERAYINWLDHAEELVALAGHTVEVTCAWSSSHDPCIVTCSDDARHKIWRIGPDLDGL 390

QY 392 -----RGLBEKPGGDK----- 402

ALIGNMENTS

hypothetical prote
hypothetical prote
beta transducin-li
WD-repeat protei
hypothetical prote
WD-repeat containi
protein T10024.21
web1 protein homol
beta transducin-li
hypothetical prote
beta-5 GTP-binding
platelet-activatin
hypothetical prote
stres protein p66
LIS-1 protein - hu

```
Db 391 SEAREAERYGTASYVREKKAFFGSSGNHKNLRLDLESTPRSLKRLMDQNERTPGSVE 450
QY 403 -----LSTVGWASQ-----KKESRRLGLVTVTSSTQTPAKAPRVK--CNPS 441
Db 451 KTTTKRSFLEMLGVAGQETEAEPQOKRAKPLESRGRLFGPSSQTA CRH1QLQISNEE 510
QY 442 NSSPSSAAACAPSCAGDPLPLPSNTPTSIKTPAKARSPINRRGSSVSPKPPSPFKMSI 501
Db 511 DASPSXKQKNSAAEDV-----SPLHKLJSTP--SHSPLSE--NVNHNVTSPPTTS---- 557
QY 502 RNWVTRTPSSSPITTPASSETKIMSPRKALIPVQKSSQAE--ACSESNNRVKRRLLDSSC 559
Db 558 ----AAAAAADAALNPPPIAALYSTSNL--PNYVLDGAPHLGIMSPRKAKEKVD--W 611
QY 560 LESVKQKCVKSCNVCVTELDGQVENLHLDLCLAGNODLSKDSLGPTKSKIEGAGTSTS 619
Db 612 LTNIRKQKLMGRAHVTLSEKI-----SEEQADVLA---SPRLQSLRQSEC 655
QY 620 EPPSPISP-----YASESCGTLPLPLPCGEGSEMVGKENSPPENKNWLLAAAKRAEN 674
Db 656 SPRIHASPRRIRISHTDGGGT-----PAGSSSH--SHSQSQPKTPT-----SSRNSET 702
QY 675 P-----SPRSPSQTPNSRRQSGKTLPSVITI-----TPSSMR 707
Db 703 TLLRPFISQSSVPAEETTTNAAPSSSDPHPAVTAAPLPLSMR 749

RESULT 2
T24284
hypotheoretical protein T01C3.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T24284
R:Wild, A.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z19869
A:Accession: T24284
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-727 <WIL>
A:Cross-references: UNIPROT:Q22059; UNIPARC:UPI0000081E3P; EMBL:Z78413; PIDN:CA01663.1;
A:Experimental source: clone T01C3
C:Genetics:
A:Gene: CBSP:T01C3.1
A:Map position: 5
A:Introns: 39/3; 88/1; 153/1; 306/1; 553/2; 654/2

Query Match 9.6%; Score 370.5; DB 2; Length 727;
Best Local Similarity 20.9%; Pred. No. 7.2e-15;
Matches 173; Conservative 127; Mismatches 279; Indels 249; Gaps 35;

QY 22 QVP-----LQSLLTGYQCNGNDHTSYGTVGPVPPFGCTSSAPNMEHVLAVANEEGF 75
Db 15 RYPSIYEELSKHYSMAGETDDHTW-----VTARFSPHLNQEHLVYMGDDPN 65
QY 76 VRLYNTESQSFRRCKPKE-----WMAHNAVFDLAWVP--GELKLVTAAGDQTAKEFWDVKA 129
Db 66 IGIFDVRKFDQSRVPLEERQLYFFPAHDGAIMDVGVGPQKESQIVSISGDSITRCWDLNQ 125
QY 130 GELIGTCK--GHQCSLKSVAESKPEKAVFCTGGRDGNIMVMDTRCNCKDGFVRVQVQIS 186
Db 126 STLDRKSQVFFGHEGSVRSICFAPDPNPFVTVGGDRDFQVKIMDMRVSTVTKMGEEDCRMAT 185
QY 187 GAHNTSDKQTPSK-----PKKQNSKGLAPSVDFOQSVTVLFDQENTLVSAGA--VDGI 239
Db 186 ITYKTAHPK--PSKVLTSIGTPKSAKAKTIE-----GYKVTSLVFLDEHHVASASENADSG 239
QY 240 IKWDLRKNYATVROBPIASKSFLYPGSSSTRKLGYSLSILDSTGSLTFLFANCTDDNIYMN 299
Db 240 IRVWDIRKPTRNGEOP--ARILKVTPTNNKKSQVGTCLTLDRGNRLFASCTDSSIFEYS 297
QY 300 MTGLKTSPTSP-----VAIFNGH 314
```

```
Db 298 VPSESVSPKFLVSHQLRKSQTNKSIITLWHSMSCASTTTLPTHVLDOYTKNLELSVNSYTCGA 357
QY 315 QNSTFYVKSLSPPDQFLVSGSSDEAAIYWKVSTPWQ-----PPTVL 356
Db 358 TIHNFYTVACSPVDVIACGSEDSERAVVMDLQDOYMNDRKLPDDIDKRRTKLPREFSC 417
QY 357 LHSQVETVSVWCPSDFTKIATCDSDNTLKTWRLNRLGLEEKPGDKLSVGH--ASQKKKE 415
Db 418 DGHKQVILNVG--SRGTYFMSC--DEGGVRIW-----SEPRNRCWTWLNDEDDT 464
QY 416 SRPGLVTVTSOSTPAKAPRVKCNPNSSPSAAACAPSCAGDPLPLPSNTPTPTFSIKTSPAK 475
Db 465 SYP-----TTSQELGLSEYKIKKFKILKESDEAMSCFD-----SISLSPQ 504
QY 476 -----ARSPINRRGSSVSPKPPSPFKMSIRNWVTRTPSSSPPI-----TPPASETKI 524
Db 505 RADSSGLSGSPQKNRGSKRPI-----FESPLKSICTNSPK--PLRLNRSPPRAKMSKL 554
QY 525 -MSPRKALIPVQKSSQAEACSESNNRVKRRLLDSSLESVKQKCVKSCNVCTE--LDGOV 581
Db 555 SFSPPSPLOPTN-----SNNQDLVGYRTP-----ROIRNKKKKNNPFFYNEHPTEGLP 601
QY 582 ENLHLDLCLAGNODLSKDSLGPTKSKIEGAGTISISEP----- 621
Db 602 NFVY-----DTFVKKULGESSKSDIEDSGKLSKTGOKRIEDMWQTKGENVATVT 651
QY 622 ----PSPISPVASESCGTLPLPLPCGEGSEMVGKENSPPENKNWLLAAAKRAENPSP 677
Db 652 RARLPS-VSEFGESAC-----SKVITEDER-----IALHSPRKL---VL 686
QY 678 RSPSSQTP---NSRRQSGKTLPSVITI-----SSMRKICTYFHR 715
Db 687 KSSSTQSPCPNSK-----PKFIPMTPRKPMDKRTPSRNLLHYFKK 727

RESULT 3
T48835
lethal(2)identicleless related protein [imported] - Neurospora crassa
A:Alternate names: protein 68B2.190
C:Species: Neurospora crassa
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: T48835
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Faritmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z24541
A:Accession: T48835
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-810 <SCH>
A:Cross-references: UNIPROT:Q9P6V7; UNIPARC:UPI000006B8C6; EMBL:ALJ53821; GSPDB:GN00112;
A:Experimental source: cosmid contig 68B2; strain 74
C:Genetics:
A:Gene: NCSP:68B2.190
A:Map position: 2
A:Introns: 236/1

Query Match 8.8%; Score 340; DB 2; Length 810;
Best Local Similarity 23.1%; Pred. No. 5.9e-13;
Matches 123; Conservative 73; Mismatches 149; Indels 188; Gaps 20;

QY 49 PVPFGCTGFSSAPNMEHVLAVANESEGFVRLYNTESQSFRRCKCFKEMAHNAVFDLAWVP 108
Db 281 PTP-----FADHPKAKNLVAVGDEEGYRLDQTQNEF--SQINMGFOAHGNAIIDAFSD 334
QY 109 GELKLVTAAGDQTAKEFWDVKAAGELICTGKHOCSLKSVAF--SKPEKAVFCTGGRDGNIM 166
Db 335 DDLATATAGDQTKVLDFTVQTPISILGHHTASLAKQVRFQPGRGANSVLATTSRDGSVQ 394
QY 167 VMDTFC-----NKKDGFYRQ-----VNOISGAHNTSDKQTPSKPKKKQNS 206
Db 395 IWDLCRGPGVQDVAIRSEAGLHARLPRTNPGCVNSIYDAHS---RIPQ--KOARNL 448
QY 207 KGLASVDF-----QQSVTVVLVFDQ---ENTLVSAGAVDGIIVKWLDRKNYATV 252
```

```

Qy 319 FYVKS-SLSPDDQFLVSGSDEAAYTKWSTPWPQPTVLGLHGSOEYTSVCMWCPSPDFTKIA 377
Db 1286 -WNVSVAFNPDGMLASGSDQTVRLWEISSKCLHT-FQHTSWSSVTTFSP-DGTMLA 1342

Qy 378 TCSDDNTLKIWRNRG 393
Db 1343 SGSDQTVRLWSISGG 1358

RESULT 5
T37884
transcription factor CDC10 target protein - fission yeast (Schizosaccharomyces
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T37884
R:Skelton, J.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August 1997
A:Reference number: Z21751
A:Accession: T37884
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-490 <SKE>
A:Cross-references: UNIPROT:Q10990; UNIPARC:UPI000013AA75; EMBL:Z98597; PIDN:CA
A:Experimental source: strain 972h-; cosmid c17H9
C:Genetics:
A:Gene: cdt2; SPDB:SPAC17H9.19c
A:Map position: 1

Query Match 7.4%; Score 286.5; DB 2; Length 490;
Best Local Similarity 26.4%; Presd. No. 5.5e-10;
Matches 96; Conservative 67; Mismatches 144; Indels 57; Gaps 14;

Qy 52 PFGCTSSAPNMEHVLAVANEEGFVRLY-----NTESQSRKKCFKEMMAHNAVFD 103
Db 130 PFCUGFA-----NNESLAVCTTETGALELFDLSRFDYDRQNEENQPSARR-IHGLWLAHNNAIFS 185

Qy 104 LAWVPGEKLIVTAAGDQTAKFWDVKAGELI-----GTCKGHQCSLKSAVFSKFEKAVFC 157
Db 186 VNFSDKDSLLATSSGQDTSKVPFLDSTQOCITRLGRRGVGDGVHSHVQVNFVNCNDSPTNLY 245

Qy 158 TGERDGNIMVWDFRCN--KDG--FYRQVNIQSAHNTSKDQTFSPKPKKQNSGLAPSV 213
Db 246 SCERDGSIIIFWDMRTTGITIDGSHFQKPVLRIRKAHENSGRDC----- 288

```

289 -----SITSATWLPOSTQVSISSCSANSAKLMDLRTVHTV-RLPAAATTPEL---TTSKRD 341

272 LGYSLLILDSTGTLFANCTDDNIYNFMNMTGLKTSPIVAIFNGH--QNSTFYVAKSSLSPPDD 329

342 FGVNVCSTPDGRIYAASRDSIIYEYSSRHLSNGFCKTKYDPLRLISSFYVKLACSPDG 401

330 QFLVSG-----SSDEAAAIWKVSTPWPQPTVLL-CHSQEVTSCVWCPSDFTKIATCSDDNT 384

402 ATLACGGGVQDKTSGVVVFPFTTENCSSAMLTGGHTKDVTAIDW--SSEQLASISDDGS 459

385 LKIW 398

460 VRVM 463

RESULT 6

AG1889

WD-40 repeat protein (imported) - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120

A:Title: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004

C:Accession: AG1889

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; DNA Res. 8, 205-213, 2001

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AG1889
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-934 <KUR>
A:Cross-references: UNIPROT:O8YZ23; UNIPARC:UPI00000CE10; GB:BA000019; PIDN:BA072622.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all0664

Query Match 7.2%; Score 279.5; DB 2; Length 934;
Best Local Similarity 22.4%; Pred. No. 3.3e-09;
Matches 95; Conservative 67; Mismatches 170; Indels 93; Gaps 13;

QY 19 WSSQVPLQSLGTYQCQSGNDEHTSYGETGVPVPPGCTFSSAPNMEHVLAVANEGFVRL 78
DB 436 WREGKLLHLEGHKDKVN-----SITES--PDQG-LIATVGVNDTKL 476
QY 79 YNTESQSFKKCFKEMAHNAVFDLAWVPGLKLVTAAGDQTAKFWDVKAGELIGTCCKG 138
DB 477 WNLDDGKELR-----TFRGHQDMWVSFSPDGKQIATASGDRTVKLWLSLDGKEL-OTLRG 530
QY 139 HQCSLKSVAFSKPEKAVFTCGRDGNIMVMDTRCNKKDGFYRQVQISGAHNTSDKQTPS 198
DB 531 HQGVNSVTFSPDGKLI-ATASGDRTVKLWNSKGQELETLYGHTD----- 574
QY 199 KFKKQNSKGLAPSDVDFQOSVTVVLPQDENTLVSAGAVDGIIVKWDLRKNYTAIROEPIA 258
DB 575 -----AVNSVAFSPDGTISIATAGNDKTAIKWLN-----S 604
QY 259 SKSFLYPGSSTRKLGYSLLDSTGLTFANCTDDNIYMFNMTGLKTSVPAIFNGHONST 318
DB 605 PMSIIVRGHEDEVF---DLVFPNGKYIATASWDKTAUKWSIVGDKLQELRTFNGHQGRV 661
QY 319 FYVKSLSLSDQDFLVSGSSDEAAYIWKVSTPMPQPTVLLGHQSQVTSVCWCPDSDFTKIAT 378
DB 662 --NKLFSFDPGKYIATTSWDKTAUKWLDGTLQ--KTLTGKDTWVSVNFSF-DGQLIAT 716
QY 379 CSDDNTLKIWLNRGLLE-----KPGDKLSTVGNASQKKESRPGLVTVT 424
DB 717 ASDEKTVKLWNRDGLLTLPRQSSVNSAVFSPDGKLIATAGWDKTVKINSIDGRLOKT 776
QY 425 SSQST 429
DB 777 LTGHT 781

RESULT 7
AH2195
hypoetical protein alr3119 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AH2195
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH2195
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-676 <KUR>
A:Cross-references: UNIPROT:O8YSG6; UNIPARC:UPI00000CE698; GB:BA000019; PIDN:BA074818.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr3119

Query Match 7.0%; Score 271; DB 2; Length 676;
Best Local Similarity 26.8%; Pred. No. 7.1e-09;
Matches 80; Conservative 48; Mismatches 115; Indels 56; Gaps 9;

QY 97 HNAVFDLAWPGLKLVTAAGDQTAKFWDVKAGELIGTCCKHQSLSKVAFSKPEKAVF 156

DB 390 HASDVNSVAFSPNGEFLASGDDKTIKVMNLKNQKIHTLPHSGVWVAIAFSPDGKTLA 449
QY 157 CTGGRDGNIMVMDTRCNKKDGFYRQVQISGAHNTSDKQTSKPKKQNSKGLAPSDVDF 216
DB 450 STGA-DKTIKWLNLATGK-----EIRHLK----- 475
QY 217 QSVTVVLPQDENTLVSAGAVDGIIVKWDLRKNYTAIROEPIASKSF--LYPGSSSTRKLG 274
DB 476 QGVASVAFSPDGKTLASGSLDKTIKWN-----PATGKEIRTLQEHSS-----GV 520
QY 275 SSLILDSTGLTFANCTDDNIYMFNMTGLKTSVPAIFNGHONSTFYVKSLSLSDQDFLVS 334
DB 521 ANVAFSPDGKTLASGSDKTIKWLNLATSKV--IHTLKGH--SDLVMSVAFNSDSQTLAS 576
QY 335 GSSDEAAYIWKVSTPMPQPTVLLGHQSQVTSVCWCPDSFTKIATCSDNTLKIWLNRG 393
DB 577 GSKDKTIKWLNLST-GKTIRTLRSHSDKVNVAVPRDSTVLASGSDNTTIKWLNLATG 634

RESULT 8
AH2154
WD-repeat protein [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AH2154
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AH2154
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1189 <KUR>
A:Cross-references: UNIPROT:O8YTD1; UNIPARC:UPI00000CE56F; GB:BA000019; PIDN:BA074490.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr2791

Query Match 6.9%; Score 268; DB 2; Length 1189;
Best Local Similarity 21.9%; Pred. No. 2.2e-08;
Matches 129; Conservative 94; Mismatches 233; Indels 132; Gaps 24;

QY 51 PPGCTFSSAPNMEHVLAVANEGFVRLYNTESQSFRRKCKPEKWA---HNAVFDLAWV 107
DB 703 PVAVHFSFGKGNM---IATASRDGTAKVWLDG-----KELSLGCHKWVWVNFPS 751
QY 108 PGLKLVTAAGDQTAKFWDVKAGELIGTCCKHQSLSKVAFSKPEKAVFTCGRDGNIMV 167
DB 752 EDGKMLVTASDRDKTAIKWLDQKEL-ATLRGHSDFTVASAVFSR-DGQIATASSDKTVRL 809
QY 168 WDTRCNKKDGFYRQVQISGAHNTSDKQTSKPKKQNSKGLAPSDVDFQOSVTVVLPQD 227
DB 810 WNRKGEELQVFWGHTDVAWG-----VNLSDKG 836
QY 228 NTLVSAGAVDGIIVKWDLRKNYTAIROEPIASKSFLYPGSSSTRKLGYS--SLILDSTGST 285
DB 837 KLLVSSGE-DGTVRLWNNGEAGKFQ---SLSFNLGEAAAGTISFSPDGKILGTGTY 891
QY 286 LFANCTDDNIYMFNMTGLKTSVPAIFNGHONSTFYVKS-SLSPDQDFLVSGSSDEAAYIW 344
DB 892 TWAK-----LWNHQG---QELVTLNGHSDT---LRLQFSPDQIATASDRDKTVKLM 938
QY 345 KYSTPMPQPTVLLGHQSQVTSVCWCPDSFTKIATCSDNTLKIWLNL----- 391
DB 939 NLN--GKERATLHGHOADVRSATPSP-DSKTIASASWDTTVKWLNLNGREIMTLRHOAG 995
QY 392 -RGLEKEFGDKLSTVGNASQKKESRPGLVTVTSQSQTAPKAPVKCNPSNSSFSSAAC 450
DB 996 VRNVFSFSDQIATASDESDGTAUKWNRQGLVTVT-----LKGHOAGIQAVSFSDQVI 1049
QY 451 APSCAGDLPLPSNTPTFSIKTSPAK-ARSPINRRGVSVSVPKP-----PSSFKMSIRN 503

Db 1050 ATA-----SKDKTVKLNROCKELLTLTLCGRGEVNAVSFSPNRETATASEDMTVKL 1101
QY 504 WYTRTPSSSPPIPPASETKIMPRKALIPVQKSSQAEACSRNRVGRRLDSSCLESV 563
Db 1102 WNLK-----GQMQTLTSLDAGVKSVSFSPDGKVLASSDSLGKVTLLWNLDPDSS 1150
QY 564 KQKCV-KSCNCVTE--LDGQVENLHLDLCCLAGNQEDLSKDSLGPRTKS 608
Db 1151 PEKLLAQCNCVRDYLLNSADIKHEERVLC-----DKLGTVKA 1188

RESULT 9
AE1810
WD-40 repeat protein [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AE1810
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AE1810
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1227 <KUR>
A:Cross-references: UNIPROT:Q8Z0R1; UNIPARC:UPI000000CDB5; GB:BA000019; PIDN:BAB77553.1
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr0029

Query Match 6.7%; Score 261.5; DB 2; Length 1227;
Best Local Similarity 23.7%; Pred. No. 5.7e-08;
Matches 102; Conservative 65; Mismatches 166; Indels 97; Gaps 16;

QY 17 NGSSQYVQLSLTGYOCSSGNDHTSYGETGVPPVPGCTFSAPNMEHVLAVANESEGV 76
Db 792 SGEDSTVRLMDVKTG-QCWQIFGHSKKVYSVSPDGQT-----LASCGBDRSI 840
QY 77 RLYNTESQSPKRCFKEMMAHNAVFDLAWVPGLKLVTAAGDQTAKFMDVKAGELI--- 133
Db 841 KLWDIQ----RGECVNTLWGHSSQVWAIASFDPGRTILSCSDDTQARLWDVITGNSLIL 896
QY 134 -----GTC-----KGHQSLSKSVAFSPKEKAV 155
Db 897 RGYTRDVYSVAFSPDSQILASGRDDVTYIGLWNLKTGECHPLRGHQRIRSVAFHPDGK-I 955
QY 156 FCTGGRDGNIMWDTNCRNKKDGFYRQVNOISGAHN-----TSDKOTPSKPKK----- 202
Db 956 LASGSADNTIKLWDISDTHSKYIR---TLTGHNNWVTVVFPDQKHTLASSSEDRIRL 1012
QY 203 --KONSKGLAPSVDFQOSVTVLWLFQDENTLVLSAGAVDGIKWDLRKNYAYRQEPITASK 260
Db 1013 WDKTGDCLQKLKGSHWVTVAFSPDGRILASGSADSEIKWD-----VASG 1060
QY 261 SFYPSGSSTRKGLYSSILSDSTGLFANCTDD-NIYFMNMTGLKTSF-VAIFNGHQNST 318
Db 1061 KCLQTLTDPQGMWIS--VAFSLDGLTLLASASEDQTVKLN--LKTGECVHTLKGEKQV 1115
QY 319 FYVKSLSPPDQFLVSGSSDEAAYIKVSTPWQPTVLLGHSQEVTSVCWCPSPDFTKIAT 378
Db 1116 YSV--AFSPNGQIAAGSSEDTTVKLDWIDISTGSCVDTLKKGHTAAIRSVAFSP-DGRLLAS 1172
QY 379 CSDNTLTKIW 388
Db 1173 GSEDEKIQWL 1182

RESULT 10
AC1842
WD-40 repeat protein [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AC1842
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AC1842
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1747 <KUR>
A:Cross-references: UNIPROT:Q8Z0Z0; UNIPARC:UPI000000CDC1; GB:BA000019; PIDN:BAB77807.1
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all0283

Query Match 6.7%; Score 260; DB 2; Length 1747;
Best Local Similarity 27.6%; Pred. No. 1.1e-07;
Matches 108; Conservative 64; Mismatches 157; Indels 62; Gaps 21;

QY 67 LAVANESEFVRLYNTESQSF----RKKC---PKEMMAHNAVFDLAWVPGLKLVTAAGD 1119
Db 1193 LATANADYTIKLYALDTSCLIVNNLQKCIQLIKTFPGHTDIVTVVFSFSDSKTIVSSSLD 1252
QY 120 QTAKFMDVKAGELIGTCGHQSLSKSVAFSPKEKAVFCTGGRDGNIMWDTNCRNKKDGFY 179
Db 1253 KTIKLRID-GSIINTWNAHGNWNSISFSPDGKMI-ASGGEENLVKLMQA----TNG-- 1304
QY 180 ROVNOISGAHN--TSDKQTP-----SKPKKQNSKG--LAPSVDFQOSVTVVLF- 224
Db 1305 HLKTLTGHERITSVKFSPDGKILASASGDKTIKFWNTDGRFLKTLIAHNNQOVNINFS 1364
QY 225 QDENTLVLSAGAVDGIKW---DLRKNYAYRQEPITASKF-----LYPGSS-----T 269
Db 1365 SDSKTLVSAGA-DSTMVKWKIDTLIKTISG-RGEQIRDVTFSPDNKVIASASSDKTVRI 1422
QY 270 RKLGY-----SSLSDSTGLFANCTDDNIYFMNMTGLKTSFVAIFNGHQNSTFYV 321
Db 1423 RQNLVQKSOXSNVSVFNPDKGTASAGWDGNITWQRETLAHSLSLTQKQNIIITV 1482
QY 322 KESLSPPDDQFLVSGSSDEAAYIKVSTPWQPTVLLGHSQEVTSVCWCPSPDFTKIATCSD 381
Db 1483 --SYSPDGTIATASADNTIKLWDSQTO-QLIKTLTGKDKRITLTSFHPDNQI-IASGSA 1538
QY 382 DNTLKIWRNRG--LEEKFG-GDKLSTVGWA 409
Db 1539 DKTIKIWRVNDGQLLTLTGHNDVTSVNFVS 1569

RESULT 11
T18521
beta transducin-like protein - Podospora anserina
C:Species: Podospora anserina
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18521
R:Saube, S.; Turcq, B.; Begueret, J.
Gene 162, 135-139, 1995
A:Title: A gene responsible for vegetative incompatibility in the fungus Podospora anserina
A:Reference number: Z18944; MUID:96009891; PMID:7557402
A:Accession: T18521
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1356 <SAU>
A:Cross-references: UNIPROT:Q00808; UNIPARC:UPI000012C621; EMBL:L28125; NID:G607002; PID
C:Genetics:
A:Gene: het-el
A:Introns: 761/3

Query Match 6.7%; Score 259.5; DB 2; Length 1356;
Best Local Similarity 25.1%; Pred. No. 8.6e-08;
Matches 94; Conservative 57; Mismatches 134; Indels 89; Gaps 14;

QY 35 SGNDHTSGETGVPPVPPGCTFSSAPNMEHVLAVANBEGFVRLVNTESQSRKCKFKEM 94
DB 942 SGSDHT-----IKINDAASGT-----CTQTL 963
QY 95 MAHNAVFDLAWVPGELKLVTAAGDQAKFVDVKGAGELIGTCKGHQCSLKSVAFSKFEKA 154
DB 964 EHGSGSVLVAFPDQORVASGDKTIKIWDATSGTCTQTLLEGHGGVSWVAFSPDQGR 1023
QY 155 VFTCTGGRDGNIMWDFTRCNKKDGFYRQVNOISGAHNSTDKQTPSPKKQNSKGLAPSV- 213
DB 1024 V-ASGSDDKTIKIWDI---ASGTCTQTLLEGHGGVQSVFSPDQORVASGDDHTIKIW 1078
QY 214 -----DFQSVTVVLFDENTLVSAAGVDGIIKWDLRKNYTAIRQEPFASKS 261
DB 1079 DAVSGTCTQTLLEGHGGVSWVAFSPDQORVASGSDGTIKINDAAS----- 1124
QY 262 FLYPGSSTRKL-GYS----SLILDSTGTLFANCTDDNIYMFN-WTGLKTSVPAIFNGHO 315
DB 1125 ----GTCTQTLLEGHGGVSWVAFSPDQORVASGSDGTIKINDAASGTCTQTL---EGHG 1177
QY 316 NSTFYVKS-SLSPDDQFLVSGSDDEAAYIKVSTPMQPTVLLGHSGEYTSVCMCPSPDFT 374
DB 1178 G---WQSVAFSPDQORVASGSDDKTIKIWDATSGTCTQTLLEGHGGVQSVAFSP-DGQ 1232
QY 375 KIATCSDDNTLKIW 388
DB 1233 RVASGSSDNTIKIW 1246
RESULT 12
AI2155
WD-repeat protein [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AI2155
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; PMID:21595285; PMID:11759840
A:Accession: AI2155
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1258 <KUR>
A:Cross-references: UNIPROT:Q8YTC2; UNIPARC:UPI000013C036; GB:BA000019; PIDN:BA074499.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr2800
Query Match 6.7%; Score 258.5; DB 2; Length 1258;
Best Local Similarity 22.6%; Pred. No. 9e-08;
Matches 93; Conservative 72; Mismatches 147; Indels 99; Gaps 16;
QY 67 LAVANEEGFVRLVNTESQSRKCKFKEMAHNAVFDLAWVPGELKLVTAAGDQAKFWD 126
DB 825 LASGSGDRTIKIWNHYT---GELKTYIGTNSVSYAIASPSDKILVSGSDRTIKLWD 880
QY 127 VRAGELIGTCKGHQCSLKSVAFSKFEKAVFC----- 157
DB 881 CQTHICIKTLHGHTNEVCSVAFSPDQTLACVSLDOSVRLMNCRTGQCCLKWYGNTDWAL 940
QY 158 -----TGGRDGNIMWDTNCKK-DGFYRQVNOISGAHNSTDKQTPSPKPKKKQ 204
DB 941 PVAFSPDRQILASGNDKTVKLWDQTKYISLEHTDFYIGAFSPDSQTLASASTDS 1000
QY 205 NSK-GLAPSVDFQ-----QSVTVVLFDENTLVSAAGVDGIIKWDLR-----KNYTA 251
DB 1001 SVRLMWNISTGQCQFILLEHTDWYAVVFPQGRKIATGSADCTVKLWNISTGQCLKLTSE 1060
QY 252 YRQ-----EPIASKSF-----LYPGSSTRKLG-----YSSLILDSTGSL 286

DB 1061 HSDKILGNMSPDQGLLASASADQSVRLWDCTGRCVGLIRGHSNRVYSATP--SPNGEI 1118
QY 287 FANC-TDNNIYMFNMTG---LKTSEVAIFNGHONSTFFVKSSLSDDQFLVSGSDDEAY 342
DB 1119 IATCSTDQTVKIDWQOGKCLKT-----LTGHTNWVFDI---AFSPDGKILASASHDQTVR 1171
QY 343 IWKYSTPMQPTVLLGHSGEYTSVCMCPSPDFTKIATCSDDNTLKIWLNRG 393
DB 1172 IWDVNT-GKCHHICIGHLHVSVAFPSP-DGEVVASGSDQTVRIWNVKTG 1220
RESULT 13
AE1866
WD-40 repeat protein [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: AE1866
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; PMID:21595285; PMID:11759840
A:Accession: AE1866
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1708 <KUR>
A:Cross-references: UNIPROT:Q8Y2I2; UNIPARC:UPI00000CDD76; GB:BA000019; PIDN:BA072436.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all0478
Query Match 6.7%; Score 258.5; DB 2; Length 1708;
Best Local Similarity 23.7%; Pred. No. 1.3e-07;
Matches 108; Conservative 61; Mismatches 168; Indels 119; Gaps 19;
QY 19 WSSQYPLQSLTGYQCSGNDHSTSGETGVPVPPGCTFSSAPNMEHVLAVANBEGFVRL 78
DB 1215 WSREGKLTLSGH-----ND-----AVLGIAWTPDQGT-----LASVGADKNIKL 1255
QY 79 YNTESQSRKCKFKEMAHNAVFDLAWVPGELKLVTAAGDQAKFWDVKA----- 129
DB 1256 WNRDG-----KLLKTWQGHDDAILGVAMSPKGTIATASFDTIKLWNRQGNLLKTLSGH 1310
QY 130 -----GELTCTCKGHQCSLKSVAFSKFEKAVFT 158
DB 1311 TAGVTAVTAVTSPNGETIGSASIDATLKLWSPQGLLGLTLKGHSNWSVSPSP-DGRIFAS 1369
QY 159 GGRDGNIMWD-----TRCNKKDGFYRQVNOISGAHNSTDKQTPSKPKKQNSKGLAPSV 213
DB 1370 GSROKTVTLRWDEVLRLNPKGDG---NDWVTSISFSSDGETLAAASRDQTVKILSRHG 1425
QY 214 DFQ-----SVTVVLFDENTLVSAAGVDGIIKWDLRKNYTAIRQEPFASKSFLYPG 266
DB 1426 KLLNTFKGHTGSIWGVAMSPNRQMIAASAKDQTVKLW-----QDG 1466
QY 267 SSTRKL-GYSSLIID---STGTLFANCTDDNIYMFNMTG---LKTSPVAIFNGHONSTF 319
DB 1467 KILHTLQHQDAVLAVAWSSDSQVIASAGKQIVKWSQGGQLLHT-----LQGHDTAVN 1521
QY 320 YVKSSLSDDQFLVSGSDDEAAYIKVSTPMQPTVLLGHSGEYTSVCMCPSPDFTKIATC 379
DB 1522 WY--SFSPDGKILASVSDTTVCLW--SRDQGLLHTLKEHSRRVGVAMSP-DGQILASA 1576
QY 380 SDNTLTKIWLNRGLEEK-PG-GDKLSTVGVASQKK 413
DB 1577 SIDGTVKLWNRDGSLSRNLPGDGDSPFISVSPSPDGK 1612
RESULT 14
C84870
probable splicing factor [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 05-Oct-2004
C;Accession: C84870
R;Lin. X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, A.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: C84870
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-343 <STO>
A;Cross-references: UNIPROT:O22826; UNIPARC:UPI00000AA2EC; GB:AB002093; NID:g2281093; PI
C;Genetics:
A;Gene: At2g43770
A;Map position: 2

Query Match 6.6%; Score 256.5; DB 2; Length 343;
Best Local Similarity 25.6%; Pred. No. 2.3e-08;
Matches 79; Conservative 44; Mismatches 117; Indels 69; Gaps 9;

QY 89 KCFKEMAHNAVFDLAWPGEIKLYTAAGDQAKFWDVYKAGELIGTCCKHQCSLKSVAF 148
DB 87 KNFWLKGHNALDLHMTSDGSIQVSASPDKTVRAWDVETGKQIKMAEHSFVNSCCP 146
QY 149 SKPEKAVFCTGGRDGNIWMDTRCNKKGDFYRQVNOISGAHNTSDKQTPSKPKKQNSKG 208
DB 147 TRGGPPLIISGDDGTAKLWDMR-----QRG 172
QY 209 LAPSVDFPQQSVTVVLFQDENTLVLSAGAVDGIKIVDLRNKNTAYROEPIASKSFLYPGSS 268
DB 173 AIQTFFPKYQITAVSFSDAAADKIFTGGVNDVVKVMDLR-----GEA 214
QY 269 TRKL-GYSSLI-----LDSTGSTIFANCTDNIYFWNT--GLKTSVPAIFNGHNS--TF 319
DB 215 TMTLEGHQDTITGMSLSPDGSYLLTNGMDNKLGVDMRPYAPQRCVKIPFGHQHFEKN 274
QY 320 YVKSSLSPPDQFLVSSGSDAAAYIKVVS---TPQPPVLGHSEVTSVCWCPSPDFTKI 376
DB 275 LLKCSWSPDGTKTAGSSDRMWHIWDTTSRRTYKLP----GHGTSVNECVFHPTE-PII 329
QY 377 ATCSDNTL 385
DB 330 GSCSSDKNI 338

RESULT 15
S19487
hypothenical protein YCR072c - Yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
C;Date: 31-Mar-1992 #sequence_revision 04-Dec-1992 #text_change 05-Oct-2004
C;Accession: S19487; S26657
R;Ballesta, J.P.G.; Franco, L.; Hoenicka, J.; Jimenez, A.; Remacha, M.; Sanz, E.
submitted to the Protein Sequence Database, March 1992
A;Reference number: S19486
A;Accession: S19487
A;Molecule type: DNA
A;Residues: 1-484, 'I', 486, 'TKL', 490, 492-515 <BAL1>
A;Cross-references: UNIPROT:P25382; UNIPARC:UPI0000143BE25; EMBL:X59720; MIPS:YCR072c
A;Note: this sequence has been revised in reference S26657
R;Ballesta, J.P.G.; Franco, L.; Hoenicka, J.; Jimenez, A.; Remacha, M.; Sanz, E.
submitted to the Protein Sequence Database, October 1992
A;Reference number: S26587
A;Accession: S26657
A;Molecule type: DNA
A;Residues: 481-503 <BAL2>
A;Cross-references: UNIPARC:UPI000017A4C5; EMBL:X59720; MIPS:YCR072c
C;Genetics:
A;Cross-references: SGD:S0000668
A;Map position: 3R
F;182-215/Domain: WD repeat homology <WD2>
F;225-265/Domain: WD repeat homology <WD3>
F;274-306/Domain: WD repeat homology <WD4>

F;398-431/Domain: WD repeat homology <WD5>
F;440-473/Domain: WD repeat homology <WD6>

Query Match 6.6%; Score 256; DB 2; Length 515;
Best Local Similarity 24.5%; Pred. No. 4.1e-08;
Matches 118; Conservative 64; Mismatches 192; Indels 108; Gaps 24;

QY 13 GVLRLNGWSSQVPLQSLTGYQCSGNDEHTSYGETGVVP-PFGCTFFSSAPNMEHVLAV- 69
DB 47 GALRVPGAISEKQLEELN--QLNGTSDD-----PVPYTFSCITQKKASDPVKTD 96
QY 70 -----ANEGFVRLYNTESSQSRKCKFEKMA---HWNVAVFDLAWVP-GE 110
DB 97 ITDNLVSSLLIKPGYNSTEDQITLLYTPRAVFKVPVTRSSSAIAGHSTILCSAFAPHTS 156
QY 111 LKLVTAAGDQAKFWDVYKAGELIGTCCKHQCSLKSVAFSKFKAFCVCTGGRGNIWMDT 170
DB 157 SRMVTGAGDNTARIWDCDQTPMHTLKGHNWVLCVSWSP-DGEVIATGSMONTIRLWDP 215
QY 171 RCNKKDGFYRQVNOISGAHN---TSDKQTP-----SKPKKQNSKGLAPSV----- 213
DB 216 KSGQCLG-----DALRG-HSKWITSLSWEPIHLVKGSKPRLASSSKDGTIKIWDTVSRV 269
QY 214 -----DFOQSVTVVLFQDENTLVLSAGAVDGIKIVMDLR-----KNYTAIROEPIA 258
DB 270 CQYTWSGHTNSVCVKWGGQGLYS--GSHDRTVRVMDINSQGRGINILKSHAHVNVNLSL 328
QY 259 SKSFLY-----PGSSTRKL--GYSSLLDSTGTLFANCTDNIYMNMTG 302
DB 329 STDYALRIGAFDHTGKPKSTPEEAQKALENYEKICKNGNSEMMVTASDDYTMFLMNP 388
QY 303 LK-TSPVAIFNGHONSTFYVKSLSLPPDQFLVSSGSDAAAYIKW-----VSTPMQPTV 355
DB 389 LKSTFPIARMTGHQKLVNHV--AFSPDGRYIIVSASFDSIKLWDGRDGRGKEIST----- 439
QY 356 LIHGHSQEVTSVCWCPSPDFTKIATCSDDNTLKIW--RLNRLGLEKPG-GDKLSTVGVWASQK 412
DB 440 FRGHIAVYQVAV--SSDCRLLVSCSKDTTLKVDVTRKLSVDLPGLHKDEVTVVDWSVDG 498
QY 413 KK 414
DB 499 KR 500

Search completed: November 8, 2006, 18:14:41
Job time : 49 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2006, 18:05:20 ; Search time 310 Seconds
(without alignments)
2178.263 Million cell updates/sec

Title: US-10-726-160-2

Perfect score: 3877

Sequence: 1 MLFNSVLRQPLQGLVLRNGWS.....TYFHRKSQDFCGPEHSTEL 730

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3877	100.0	730	2 Q9NZJ0_HUMAN	Q9nzj0 homo sapien
2	3872	99.9	730	2 Q9NMW5_HUMAN	Q9nmw5 homo sapien
3	3867	99.7	730	2 Q5V777_HUMAN	Q5vt77 homo sapien
4	3499.5	90.3	729	2 Q3TLR7_MOUSE	Q3tlr7 mus musculus
5	3496.5	90.2	729	2 Q8BW38_MOUSE	Q8bw38 m 2 days pr
6	3479.5	89.7	729	2 Q80WY2_MOUSE	Q80wy2 mus musculus
7	3330.5	85.9	732	2 Q80WY1_MOUSE	Q80wy1 mus musculus
8	2835.5	73.1	597	2 Q3TTE9_MOUSE	Q3tte9 mus musculus
9	2705	69.8	534	2 Q6PAN1_MOUSE	Q6pan1 mus musculus
10	2293	59.1	434	2 Q9NMW3_HUMAN	Q9nmw3 homo sapien
11	2206.5	56.9	720	2 Q5ZJW8_CHICK	Q5zjw8 gallus gall
12	2098.5	54.1	713	2 Q6PIW0_XENTR	Q6piw0 xenopus tro
13	2052	52.9	710	2 Q4V837_XENIA	Q4v837 xenopus lae
14	2034.5	52.5	711	2 Q6GRU3_XENIA	Q6gru3 xenopus lae
15	1630.5	42.1	647	2 Q5RH15_BRARE	Q5rh15 brachydanio
16	1622.5	41.8	647	2 Q7ZU24_BRARE	Q7zu24 brachydanio
17	1615.5	41.7	647	2 Q8JH14_BRARE	Q8jh14 brachydanio
18	1255.5	32.4	475	2 Q4SCV3_TETNG	Q4scv3 tetraodon n
19	1216	31.4	232	2 Q9NW34_HUMAN	Q9nw34 homo sapien
20	1063	27.4	216	2 Q9C276_MOUSE	Q9c276 mus musculus
21	1020	26.3	211	2 Q96SN0_HUMAN	Q96sn0 homo sapien
22	793.5	20.5	769	1 L2DPTL_DROME	L2dptl drosophila
23	783	20.2	632	2 Q7PTB3_ANOGA	Q7ptb3 anopheles g
24	652	16.8	535	2 Q94C55_ARATH	Q94c55 arabidopsis
25	652	16.8	566	2 Q9LXK5_ARATH	Q9lxk5 arabidopsis
26	575.5	14.8	555	2 Q85L86_ORYSA	Q85l86 oryza sativ
27	404	10.4	676	2 Q61FK1_CAEBR	Q61fk1 caenorhabdi
28	370.5	9.6	727	2 Q2D059_CAEEL	Q2d059 caenorhabdi
29	340	8.8	810	2 Q9P6V7_NEUCR	Q9p6v7 neurospora
30	339	8.7	712	2 Q4WZRS_ASPFU	Q4wzrs aspergillus
31	332	8.6	727	2 Q2UDF0_ASPOR	Q2udf0 aspergillus

RESULT 1												
Q9NZJ0_HUMAN PRELIMINARY; PRT; 730 AA.												
ID	Q9NZJ0	HUMAN	PRELIMINARY;	PRT;	730	AA.						
AC	Q9NZJ0;											
DT	01-OCT-2000,	integrated into UniProtKB/TrEMBL.										
DT	01-OCT-2000,	sequence version 1.										
DT	07-FEB-2006,	entry version 28.										
DE	L2DTL protein (RA-regulated nuclear matrix-associated protein)											
DE	(Retinoic acid-regulated nuclear matrix-associated protein)											
DE	(Denticleless homolog)											
GN	Name=L2DTL; Synonyms=DTL;											
OS	Homo sapiens (Human)											
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;											
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;											
OC	Homo.											
OX	NCBI_TaxID=9606;											
RN	[1]											
RP	NUCLEOTIDE SEQUENCE.											
RA	Mueller R., Ziegler B.L.;											
RL	Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.											
RN	[2]											
RP	NUCLEOTIDE SEQUENCE.											
RX	MEDLINE=21264461; PubMed=11278750; DOI=10.1074/jbc.M010802200;											
RA	Cheung W.M., Chu A.H., Chu P.W., Ip N.Y.;											
RT	"Cloning and expression of a novel nuclear matrix-associated protein											
RT	that is regulated during the retinoic acid-induced neuronal											
RT	differentiation.";											
RL	J. Biol. Chem. 276:17083-17091(2001).											
RP	NUCLEOTIDE SEQUENCE.											
RC	TISSUE=Lymph. and Testis;											
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;											
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,											
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schaefer C.F., Bhat N.K.,											
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,											
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,											
RA	Datchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,											
RA	Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,											
RA	Brownstein M.J., Usslin T.B., Toshiyuki S., Carninci P., Prange C.,											
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,											
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,											
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,											
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,											
RA	Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,											
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,											
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,											
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,											
RA	Butterfield Y.S.N., Krzyzinski M.I., Skalska U., Smalhus D.E.,											
RA	Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;											
RT	"Generation and initial analysis of more than 15,000 full-length human											
RT	and mouse cDNA sequences.";											
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).											
RN	[4]											
RP	NUCLEOTIDE SEQUENCE.											

Q55ln8 cryptococcus
Q5bbd2 aspergillus
Q3mc23 anabaena va
Q3mcv7 anabaena va
Q3hjd1 trichodesmi
Q8yrl1 anabaena sp
Q47a03 destilago ma
Q47a03 destilago ma
Q10990 schizosacch
Q7nd05 gloeobacter
Q4c9p2 crocosphaer
Q8y8g6 anabaena sp
Q4i9g2 gibberella

RC TISSUE=Lymph, and Testis;
RG NIH MGC Project;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AF195765; AAF35182.1; -; mRNA.
DR EMBL; AF345896; AAK54706.1; -; mRNA.
DR EMBL; BC033297; AAH33297.1; -; mRNA.
DR EMBL; BC033540; AAH33540.1; -; mRNA.
DR Ensembl; ENSG00000143476; Homo sapiens.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 5.
DR PRINTS; PR00320; GPROTEINBRPT.
DR ProDom; PD000018; WD40; 1.
DR SMART; SM00320; WD40; 5.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_2.
DR PROSITE; PS00882; WD_REPEATS_2; 5.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 730 AA; 79469 MW; B4A149BC62059C4F CRC64;

Query Match 100.0%; Score 3877; DB 2; Length 730;
Best Local Similarity 100.0%; Pred. No. 3.3e-206;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLFNSVLRQPLQGLVLRNGWSSQYPLQSLLTGYQCSGNDEHTSYGETGVVPVPPFGCTFSSA 60
DB 1 MLFNSVLRQPLQGLVLRNGWSSQYPLQSLLTGYQCSGNDEHTSYGETGVVPVPPFGCTFSSA 60
QY 61 PMEHVLA VANE EGFVRLYNTESQSFRKCKFKEMAHWNAVFDLAWVPGLKLVTAAGDQ 120
DB 61 PMEHVLA VANE EGFVRLYNTESQSFRKCKFKEMAHWNAVFDLAWVPGLKLVTAAGDQ 120
QY 121 TAKFWDV KAGELIGTCKGHQCSLKSVAFSKPEKAVFTCGRDGNIWMDTRCNKKDGFYR 180
DB 121 TAKFWDV KAGELIGTCKGHQCSLKSVAFSKPEKAVFTCGRDGNIWMDTRCNKKDGFYR 180
QY 181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQSQVTVVLFDENTLVSA GAVDGI I 240
DB 181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQSQVTVVLFDENTLVSA GAVDGI I 240
QY 241 KVDLRKNYTA VROEPIASKSFLYPGSSTRKLKGLYSSLLDSTGSLFANCTDDNIYMFNM 300
DB 241 KVDLRKNYTA VROEPIASKSFLYPGSSTRKLKGLYSSLLDSTGSLFANCTDDNIYMFNM 300
QY 301 TGLKTS PVAIFNGHONSTFYVKSSLSDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
DB 301 TGLKTS PVAIFNGHONSTFYVKSSLSDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
QY 361 QVTSVVCWCPSPDTKTIATCSDNTLKIWRNLRGLEKPGDKLSTVGMASQKKESRPGI 420
DB 361 QVTSVVCWCPSPDTKTIATCSDNTLKIWRNLRGLEKPGDKLSTVGMASQKKESRPGI 420
QY 421 VVTSSQSTPAKAPRVKCNPSNSSPSAACPSCAGDLPLPNTPTFSIKTSPAKARSPI 480
DB 421 VVTSSQSTPAKAPRVKCNPSNSSPSAACPSCAGDLPLPNTPTFSIKTSPAKARSPI 480
QY 481 NRRGVS SVSPKPPSPFKMSIRNWVTRTPSSSPITPPASETKIMSPKRALIPVSQKSSQ 540
DB 481 NRRGVS SVSPKPPSPFKMSIRNWVTRTPSSSPITPPASETKIMSPKRALIPVSQKSSQ 540
QY 541 AEACSSRNVRKRLDSSCLIESVKQCVKSCNCVTELDGQVENLHLDLCLAGNQEDLSK 600
DB 541 AEACSSRNVRKRLDSSCLIESVKQCVKSCNCVTELDGQVENLHLDLCLAGNQEDLSK 600
QY 601 DSLGPTKSKIEGAGTSISEPPSPISFYASESCGTLPLPLPCGEGSEVMGKENS SPENK 660
DB 601 DSLGPTKSKIEGAGTSISEPPSPISFYASESCGTLPLPLPCGEGSEVMGKENS SPENK 660
QY 661 NWLLAAAKRAKNAENPSPSPSSQTFNSRRQSGKTLPSPTITPSSMRKICTYFHRKSQED 720
DB 661 NWLLAAAKRAKNAENPSPSPSSQTFNSRRQSGKTLPSPTITPSSMRKICTYFHRKSQED 720

DB 661 NWLLAAAKRAKNAENPSPSPSSQTFNSRRQSGKTLPSPTITPSSMRKICTYFHRKSQED 720
QY 721 FCGPEHSTEL 730
DB 721 FCGPEHSTEL 730

RESULT 2
Q9NWM5 HUMAN PRELIMINARY; PRT; 730 AA.
ID Q9NWM5_HUMAN
AC Q9NWM5;
DT 01-OCT-2000, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2000, sequence version 1.
DT 07-FEB-2006, entry version 21.
DE Hypothetical protein FLJ20735.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
RA Nakajima Y., Mizuno T., Morinaga M., Ota T., Suzuki Y., Obayashi M.,
RA Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AK000742; BAA91355.1; -; mRNA.
DR Ensembl; ENSG00000143476; Homo sapiens.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 5.
DR PRINTS; PR00320; GPROTEINBRPT.
DR ProDom; PD000018; WD40; 1.
DR SMART; SM00320; WD40; 5.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_2.
DR PROSITE; PS00882; WD_REPEATS_2; 5.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 730 AA; 79479 MW; 7640815C8D05303F CRC64;

Query Match 99.9%; Score 3872; DB 2; Length 730;
Best Local Similarity 99.9%; Pred. No. 6.2e-206;
Matches 729; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLFNSVLRQPLQGLVLRNGWSSQYPLQSLLTGYQCSGNDEHTSYGETGVVPVPPFGCTFSSA 60
DB 1 MLFNSVLRQPLQGLVLRNGWSSQYPLQSLLTGYQCSGNDEHTSYGETGVVPVPPFGCTFSSA 60
QY 61 PMEHVLA VANE EGFVRLYNTESQSFRKCKFKEMAHWNAVFDLAWVPGLKLVTAAGDQ 120
DB 61 PMEHVLA VANE EGFVRLYNTESQSFRKCKFKEMAHWNAVFDLAWVPGLKLVTAAGDQ 120
QY 121 TAKFWDV KAGELIGTCKGHQCSLKSVAFSKPEKAVFTCGRDGNIWMDTRCNKKDGFYR 180
DB 121 TAKFWDV KAGELIGTCKGHQCSLKSVAFSKPEKAVFTCGRDGNIWMDTRCNKKDGFYR 180
QY 181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQSQVTVVLFDENTLVSA GAVDGI I 240
DB 181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQSQVTVVLFDENTLVSA GAVDGI I 240
QY 241 KVDLRKNYTA VROEPIASKSFLYPGSSTRKLKGLYSSLLDSTGSLFANCTDDNIYMFNM 300
DB 241 KVDLRKNYTA VROEPIASKSFLYPGSSTRKLKGLYSSLLDSTGSLFANCTDDNIYMFNM 300
QY 301 TGLKTS PVAIFNGHONSTFYVKSSLSDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
DB 301 TGLKTS PVAIFNGHONSTFYVKSSLSDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
QY 361 QVTSVVCWCPSPDTKTIATCSDNTLKIWRNLRGLEKPGDKLSTVGMASQKKESRPGI 420
DB 361 QVTSVVCWCPSPDTKTIATCSDNTLKIWRNLRGLEKPGDKLSTVGMASQKKESRPGI 420

Db 361 QEVTSVCMCPSPDTKIATCDDNTLKIWLNRGLEEKPGDKLSTVGWASQKKESRPGL 420
Qy 421 VVTSSQSTPAKAPRVKCNPSNSSPSAACAPSCAGDLPLPSNTPPTFSIKTSPAKARSPI 480
Db 421 VVTSSQSTPAKAPRVKCNPSNSSPSAACAPSCAGDLPLPSNTPPTFSIKTSPAKARSPI 480
Qy 481 NRRGVSVPSPKPPSPFKMSIRNWTPTSSSPPIPPASSETKIMSPRKALIPVSQKSSQ 540
Db 481 NRRGVSVPSPKPPSPFKMSIRNWTPTSSSPPIPPASSETKIMSPRKALIPVSQKSSQ 540
Qy 541 AEACSESRNRVRRLLDSSCLSVKQKCVKSCNCTELDQGVENHLHDLCLCLAGNQEDLSK 600
Db 541 AEACSESRNRVRRLLDSSCLSVKQKCVKSCNCTELDQGVENHLHDLCLCLAGNQEDLSK 600
Qy 601 DSIQPTKSSKIEGAGTSISEPPSPISPYASESGTLPPLRPGEGSEVMWGENSSPENK 660
Db 601 DSIQPTKSSKIEGAGTSISEPPSPISPYASESGTLPPLRPGEGSEVMWGENSSPENK 660
Qy 661 NWLLAMAAKKAENSPRSPSSQTPNSRRQSGKTLPSVPTITPSSMRKICTYFHRKSQED 720
Db 661 NWLLAMAAKKAENSPRSPSSQTPNSRRQSGKTLPSVPTITPSSMRKICTYFHRKSQED 720
Qy 721 FCGPEHSTEL 730
Db 721 FCGPEHSTEL 730

RESULT 3
Q5VT77 HUMAN PRELIMINARY; PRT; 730 AA.
AC Q5VT77;
DT 07-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 07-DEC-2004, sequence version 1.
DT 07-FEB-2006, entry version 13.
DE RA-regulated nuclear matrix-associated protein.
GN Name=RAMP; ORFNames=RP11-15111.1-001;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Almeida J.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Sycamore N.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
CC EMBL; AL606468; CAH70697.1; -; Genomic DNA.
DR EMBL; AC092814; CAH70697.1; JOINED; Genomic DNA.
DR EMBL; AL592297; CAH70697.1; JOINED; Genomic DNA.
DR EMBL; AL592297; CAH73803.1; -; Genomic DNA.
DR EMBL; AC092814; CAH73803.1; JOINED; Genomic DNA.
DR EMBL; AL606468; CAH73803.1; JOINED; Genomic DNA.
DR Ensembl; ENSG00000143476; Homo sapiens.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 5.
DR PRINTS; PR00320; GPROTEINBRPT.
DR ProDom; PD000018; WD40; 1.
DR SMART; SM00320; WD40; 5.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_2.
DR PROSITE; PS50082; WD_REPEATS_2; 5.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 730 AA; 79468 MW; CE8D54234D44F002 CRC64;

Query Match 99.7%; Score 3867; DB 2; Length 730;
Best Local Similarity 99.7%; Pred. No. 1.2e-205;

Matches 728; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MLFNSVLRPOLGVLRNGWSSQYPIQSLLTGVCSCGNDEHTSYGETGVVPVPPGCTFSSA 60
Db 1 MLFNSVLRPOLGVLRNGWSSQYPIQSLLTGVCSCGNDEHTSYGETGVVPVPPGCTFSSA 60
Qy 61 PNMEHLVAVANEEGFVRLYNTESQSPFRKCKFKEMAHNAVFDLAWVPGLKLVTAAGDQ 120
Db 61 PNMEHLVAVANEEGFVRLYNTESQSPFRKCKFKEMAHNAVFDLAWVPGLKLVTAAGDQ 120
Qy 121 TAKFMDVWVAGELIGTCKGHQCSLKSVAFSPKPKAVFCTGGRDGNIMWDTNCKNKGDFR 180
Db 121 TAKFMDVWVAGELIGTCKGHQCSLKSVAFSPKPKAVFCTGGRDGNIMWDTNCKNKGDFR 180
Qy 181 QYNQISGAHNTSDKOTPSKPKKQKSLGAPSDVDFQOSVTVVLFDENTLVSAGAVDGI 240
Db 181 QYNQISGAHNTSDKOTPSKPKKQKSLGAPSDVDFQOSVTVVLFDENTLVSAGAVDGI 240
Qy 241 KYWDLRKNTAYRQBPIASKSFLYPCSTRKLGYSLLDSTGSLFANCTDDNIYFMNM 300
Db 241 KYWDLRKNTAYRQBPIASKSFLYPCSTRKLGYSLLDSTGSLFANCTDDNIYFMNM 300
Qy 301 TGLKTSPIVAIFNGHONSFPYKSSLSPPDQFLVSGSSDEAAIYWKYSTPWQPTVLLGHS 360
Db 301 TGLKTSPIVAIFNGHONSFPYKSSLSPPDQFLVSGSSDEAAIYWKYSTPWQPTVLLGHS 360
Qy 361 QEVTSVCMCPSPDTKIATCDDNTLKIWLNRGLEEKPGDKLSTVGWASQKKESRPGL 420
Db 361 QEVTSVCMCPSPDTKIATCDDNTLKIWLNRGLEEKPGDKLSTVGWASQKKESRPGL 420
Qy 421 VVTSSQSTPAKAPRVKCNPSNSSPSAACAPSCAGDLPLPSNTPPTFSIKTSPAKARSPI 480
Db 421 VVTSSQSTPAKAPRVKCNPSNSSPSAACAPSCAGDLPLPSNTPPTFSIKTSPAKARSPI 480
Qy 481 NRRGVSVPSPKPPSPFKMSIRNWTPTSSSPPIPPASSETKIMSPRKALIPVSQKSSQ 540
Db 481 NRRGVSVPSPKPPSPFKMSIRNWTPTSSSPPIPPASSETKIMSPRKALIPVSQKSSQ 540
Qy 541 AEACSESRNRVRRLLDSSCLSVKQKCVKSCNCTELDQGVENHLHDLCLCLAGNQEDLSK 600
Db 541 AEACSESRNRVRRLLDSSCLSVKQKCVKSCNCTELDQGVENHLHDLCLCLAGNQEDLSK 600
Qy 601 DSIQPTKSSKIEGAGTSISEPPSPISPYASESGTLPPLRPGEGSEVMWGENSSPENK 660
Db 601 DSIQPTKSSKIEGAGTSISEPPSPISPYASESGTLPPLRPGEGSEVMWGENSSPENK 660
Qy 661 NWLLAMAAKKAENSPRSPSSQTPNSRRQSGKTLPSVPTITPSSMRKICTYFHRKSQED 720
Db 661 NWLLAMAAKKAENSPRSPSSQTPNSRRQSGKTLPSVPTITPSSMRKICTYFHRKSQED 720
Qy 721 FCGPEHSTEL 730
Db 721 FCGPEHSTEL 730

RESULT 4
Q3TLR7 MOUSE PRELIMINARY; PRT; 729 AA.
ID Q3TLR7_MOUSE
AC Q3TLR7;
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2005, sequence version 1.
DT 07-FEB-2006, entry version 6.
DE Mammary gland RCB-0526 Jyg-MC(A) cDNA, RIKEN full-length enriched
DE library, clone: G830048A21 product: l22DTL protein, full insert sequence.
GN Name=Dtl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;

RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Methods Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX PubMed=16141072; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
Oyama R., Ravasi T., Lenhard D., Wells C., Kodzius R., Shimokawa K.,
Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
Crome M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
Di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
Fletcher C.P., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
Hill D., Hummel L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kello J., Kitamura H.,
Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
Matsuura H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
Mortrud M., Mulder N., Nakano N., Nakaguchi H., Ng P.,
Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavoni G., Pesole G.,
Petkovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
Sperling S., Stupka E., Sugtara K., Sultana R., Takenaka Y., Taki K.,
Tamaoka K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
Kawashina T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
Tagami H., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563(2005).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX PubMed=16141073; DOI=10.1126/science.1112009;
RG RIKEN Genome Exploration Research Group, and Genome Science Group
(Genome Network Core Team) and the FANTOM Consortium;
RT "Antisense Transcription in the Mammalian Transcriptome.";
RL Science 309:1564-1566(2005).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX MEDLINE=22354663; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
Balderelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
Kanai A., Kawaji H., Kawasawa Y., Kedziere R.M., King B.L.,
Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
Nagashima T., Numata K., Okido T., Pavan W.J., Perrega G., Pesole G.,
Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
Sandelin A., Schneider C., Semple C.A., Setou M., Shinada K.,
Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
Kawaji H., Kiyosawa H., Kondo S., Kondo S., Yamanaka I.,
Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,
Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,
Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
Gustincich S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
Fujiwara S., Inoue K., Togawa Y., Izawa K., Ohara E., Watahiki M.,
Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [8]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Mammary gland;
RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,
Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,
Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
Muramatsu M., Hayashizaki Y.;
RT Submitted (APR-2004) to the ENBL/GenBank/DBJ databases.
RL -----

CC Copyrighted by the Uniprot Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL: AK166351; BAE38725.1; -, mRNA.
DR MG1: MG11924093; DCL.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 5.
DR PRINTS: PR00320; GPROTEINRPT.
DR ProDom: PD000018; WD40; 1.
DR SMART: SM00320; WD40; 5.
DR ProSITE: PS00678; WD_REPEATS_1; UNKNOWN_2.
DR ProSITE: PS00082; WD_REPEATS_2; 5.
DR ProSITE: PS0294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 729 AA; 79091 MW; C0F4C515A9404DAD CRC64;

Query Match 90.3%; Score 3499.5; DB 2; Length 729;
Best Local Similarity 89.6%; Pred. No. 2.5e-185;
Matches 654; Conservative 36; Mismatches 39; Indels 1; Gaps 1;

QY 1 MLFNSVLRLQPLGVLNRNGSSQYPLQSLLTGYQCSGNDEHTSYGETGVPPPGCTFSSA 60
DB 1 MLFNSVLRLQPLGVLNRNGSSHYPLQSLLSGYQCNDEHTSYGETGVPPPGCTFCTA 60
QY 61 PNMHEVLAVANERGFVRLYNTEQSFRKCKCFKEMAHNAVFDLAWVPGELKLVTAAGDQ 120
DB 61 PNMHEVLAVANERGFVRLYNTEQSFKCTCFKEMAHNAVFDLAWVPGELKLVTAAGDQ 120
QY 121 TAKFWDVAGELGTCKGHQCSLKSVAFSKFAVCTGGRDGNIMVWDRCKNKGDFYR 180
DB 121 TAKFWDVAGELGTCKGHQCSLKSVAFSKFAVCTGGRDGNIMVWDRCKNKGDFYR 180
QY 181 QVNOISGAHNTSKDQTPSKPKKQNSGLAPSVDFFQSVTVVLFDQENTLVLSAGVDGII 240
DB 181 QVNOISGAHNTADQTPSKPKKQNSGLAPVADVSQSVTVVLFDQENTLVLSAGVDGII 240
QY 241 KWDLRKNYTAHQEPIASKSLFYPSGSTKGLGYSSILLDSTGSTLFANCTDDNIYFNM 300
DB 241 KWDLRKNYTAHQEPIASKSLFYPGTSTKGLGYSSILVDSTGSTLFANCTDDNIYFNM 300
QY 301 TGLKTSFVAIFNGHQNSTFYVKSLSLSPDDQFLVSGSDEAAYTWKYSTPMQPTVLLGHS 360
DB 301 TGLKTSFVAIFNGHQNSTFYVKSLSLSPDDQFLVSGSDEAAYTWKYSMPHPPPTVLLGHS 360
QY 361 QEVTSVWCPSDPTKATCDDNTLKIWLNRGLGEEKPGDKLSTGCVASOKKESRPLG 420
DB 361 QEVTSVWCPSDPTKATCDDNTLKIWLNRGLGEEKP-GDKHSIVGWTSSQKKKVKACP 419
QY 421 VVTSSQSTPAKAPRVKCNPSNPSGSAACAPSCAGDPLPSPNTPTFSIKTSPAKARSP 480
DB 420 VTPVPSQSTPAKAPRAKSSPSISPSAACTPCAGDPLPSPSTPTFSVKTTPATRSSV 479
QY 481 NRRGSSVSPKPPSPFQKMSIRNWVTRTPSSSPPTPPASSETKIMSPRKALIPVSQSSQ 540
DB 480 SRRGSSVSPKPLSPFQKMSLRNWVTRTPSSSPPTPPASSETKISPRKALIPVSQSSQ 539
QY 541 AEACSESRNRKRLDSSCLSVKQKVCNCVTELDGQVENLHLDLCLCLAGNQDLSK 600
DB 540 ADACSESRNRKRLDSSCLSVKQKVCNCVTELDGQAESLRDLCLCLSGAQEVLISQ 599
QY 601 DSGPTKSSKIEGAGTSISPPSPISPYASESGTLPPLPRLPCGEGSEMWKENSPPENK 660
DB 600 DSGPTKSSKIEGAGTSISPPSPISPYASEGCGPLPLPRLPCGEGSEMWKENSPPENK 659
QY 661 NWLLAAAKKAENPNRSPSSQTPNRRSGKTLPSVPTITPSSMRKICTYFHRKSQED 720
DB 660 NWLLAAAKKAENPNRSPSSQTPNRRSGKTLPSVPTITPSSMRKICTYFHRKTQDD 719
QY 721 FCSPEHSTEL 730
DB 720 FCSPEHSTEL 729

RESULT 5
Q8BW38_MOUSE PRELIMINARY; PRT; 729 AA.
AC Q8BW38_MOUSE
DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2003, sequence version 1.
DT 07-FEB-2006, entry version 26.
DE 2 days pregnant adult female ovary cDNA, RIKEN full-length enriched
DE library, clone:E330023F01 product:L2DTL PROTEIN (RA-REGULATED NUCLEAR
DE MATRIX-ASSOCIATED PROTEIN) homolog (Adult male testis cDNA, RIKEN
DE full-length enriched library, clone:4931420L14 product:L2DTL protein,
DE full insert sequence).
GN Name=Dtl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Ovary, and Testis;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Methods Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Ovary, and Testis;
RX PubMed=16141072; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
Davis M.J., Walimbe L.G., Aidinis V., Allen J.E.,
Ambesi-Imbimbato A., Apweiler R., Attalaya R.N., Bailey T.L.,
Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
Crome M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
di Bernardo C., Down T., Engstrom P., Fagioli M., Faulkner G.,
Fletcher C.P., Fukushima T., Furum M., Futaki S., Gariboldi M.,
Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
Hill D., Hummel L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
Kitano H., Kollas G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
Mottagui-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,
Nilsen R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
Shenbrot C., Sekiguchi K., Sempere C.A., Seno S., Sessa L., Sheng Y.,
Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
Sperling S., Stupka E., Sugura K., Sultana R., Takenaka Y., Taki K.,
Tammaja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
Griffith S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
Tada J., Inamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
Kawahara T., Kohama M., Kondo H., Konno H., Nakano K., Ninomiya N.,
Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome.";
RL Science 309:1559-1563(2005).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Ovary, and Testis;
RX PubMed=16141073; DOI=10.1126/science.1112009;
RG RIKEN Genome Exploration Research Group, and Genome Science Group

RG (Genome Network Core Team) and the FANTOM Consortium;
RT "Antisense Transcription in the Mammalian Transcriptome";
RL Science 309:1564-1566 (2005).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Ovary, and Testis;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01265;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Oatono N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Bruscia V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grainger S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shinada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wegner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki K., Shibata K., Shingawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573 (2002).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Ovary, and Testis;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fletschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection";
RL Nature 409:685-690 (2001).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Ovary, and Testis;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes";
RL Genome Res. 10:1617-1630 (2000).
RN [7]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Ovary, and Testis;
RX MEDLINE=20330913; PubMed=11076861; DOI=10.1101/gr.152600;
RA

RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Iehikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer";
RL Genome Res. 10:1757-1771 (2000).
RN [8]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Ovary;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akaira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [9]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
RA Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,
RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,
RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
RA Shibata K., Shiraki T., Tagami M., Tagami M., Waki K., Watahiki A.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonCommercial License
CC -----
DR EMBL; AK054412; BAC35769.1; -. mRNA.
DR EMBL; AK133177; BAE21543.1; -. mRNA.
DR Ensembl; ENSMUSG00000037474; Mus musculus.
DR MGI; MGI:1924093; Dcl.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 5.
DR PRINTS; PR00320; GPROTEINBRPT.
DR ProDom; PD000018; WD40; 1.
DR SMART; SM00320; WD40; 5.
DR PROSITE; PS00578; WD_REPEATS_1; UNKNOWN_2.
DR PROSITE; PS00882; WD_REPEATS_2; 5.
DR PROSITE; PS02994; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 729 AA; 79131 MW; C828FAFBA9929360 CRC64;
Query Match 90.2%; Score 3496.5; DB 2; Length 729;
Best Local Similarity 89.5%; Pred. No. 3.7e-185;
Matches 653; Conservative 36; Mismatches 40; Indels 1; Gaps 1;
QY 1 MLFNSVLQPQLGVLNRNGWSSQYPLQSLITGYQCSCNDEHTSYGETGVPPVPGCTFSSA 60
DB 1 MLFNSVLQPQLGVLNRNGWSSHYPLQSLISGYQCNCNDEHTSYGETGVPPVPGCTFCTA 60
QY 61 PMEHVLAVANBEGFVRLYNTESQSFRRKCKFKEMAHNNAVDFLAWVPGELKLVTAAGDQ 120
DB 61 PMEHVLAVANBEGFVRLYNTESQSKTKCFKEMAHNNAVDFLAWVPGELKLVTAAGDQ 120
QY 121 TAKFWDVKGAGELIGTCCKHQCSLKSVAFKPKFAVFCCTGGRDGNIMVMDTRCNCKDGFYR 180
DB 121 TAKFWDVKGAGELIGTCCKHQCSLKSVAFKPKFAVFCCTGGRDGNIMVMDTRCNCKDGFYR 180
QY 181 QVNOISGAHNTSDKOTPSKPKKQNSKGLAPAVDVFQSQSVTVVLFQDENTLVSAGAVDGI 240
DB 181 QVNOISGAHNTADKOTPSKPKKQNSKGLAPAVDVFQSQSVTVVLFQDENTLVSAGAVDGI 240

```
QY 241 KWDLRKNYTAHQEPIASKFLYPGSSTRKLGYSLLILDSTGSLTFANCTDDNIYFMN 300
DB 241 KWDLRKNYTAHQEPIASKFLYPGSSTRKLGYSLLILDSTGSLTFANCTDDNIYFMN 300
QY 301 TGLKTSFPAIFNGHQNSTFYVKSLSLSPDDQFLVSGSSDEAAIYKWTSTPWPQPTVLLGHS 360
DB 301 TGLKTSFPAIFNGHQNSTFYVKSLSLSPDDQFLVSGSSDEAAIYKWTSTPWPQPTVLLGHS 360
QY 361 QEVTSVCWCPSPDFTKIATCDDNTLKIWLNRGLBEKPGDGLKSTVGWASOKKESRPGL 420
DB 361 QEVTSVCWCPSPDFTKIATCDDNTLKIWLNRGLBEKPGDGLKSTVGWASOKKESRPGL 420
QY 421 VVTSSQSTPAKAPRVKCNPSNPSAACAAPSCAGDPLPSPNTPTFSIKTSPAKARSP 480
DB 421 VVTSSQSTPAKAPRVKCNPSNPSAACAAPSCAGDPLPSPNTPTFSIKTSPAKARSP 480
QY 481 NRRGVSVPSPKPPSPFMSIRNWTPTSSSPPIITPPASETKIMSPRKALIPVSKSSQ 540
DB 481 NRRGVSVPSPKPPSPFMSIRNWTPTSSSPPIITPPASETKIMSPRKALIPVSKSSQ 540
QY 541 AEACSESRNRVRRDLSSCLSVKQKVCSCNCTVLDGQVENLHLDLCLAGNQEDLSK 600
DB 541 AEACSESRNRVRRDLSSCLSVKQKVCSCNCTVLDGQVENLHLDLCLAGNQEDLSK 600
QY 601 DSIQPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLRPLRPGEGSEMVKGKENSSENK 660
DB 601 DSIQPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLRPLRPGEGSEMVKGKENSSENK 660
QY 661 NWLLAAKRAKKAENSPRSPSSQTPNSRRSQSGTKLSPVITPSSMRKICTYFHRKSQED 720
DB 661 NWLLAAKRAKKAENSPRSPSSQTPNSRRSQSGTKLSPVITPSSMRKICTYFHRKSQED 720
QY 721 FCPSEHSTEL 730
DB 721 FCPSEHSTEL 729

RESULT 6
Q80WY2 MOUSE
ID Q80WY2_MOUSE PRELIMINARY; PRT; 729 AA.
AC Q80WY2;
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 21.
DE Retinoic-acid regulated nuclear matrix-associated protein.
GN Name=2810047L02Rik; Synonyms=Dtl, ramp;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BALB/c;
RA Uenaka A., Hirano Y., Hata H., Sanda W., Aji T., Tanaka M., Ono T.,
RA Skipper J., Shimizu K., Nakayama E.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AB095735; BAC76404.1; -; mRNA.
DR Ensembl; ENSMUSG00000037474; Mus musculus.
DR MGI; MGI:1924093; 2810047L02Rik.
DR MGI; MGI:1924093; Dtl.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 5.
DR PRINTS; PR00320; GPROTEINRPT.
DR ProDom; PD000018; WD40; 1.
DR SMART; SM00320; WD40; 5.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_2.
DR PROSITE; PS50082; WD_REPEATS_2; 5.
```

```
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 729 AA; 79161 MW; 81DAF4BACD5F38B1 CRC64;

Query Match      89.7%; Score 3479.5; DB 2; Length 729;
Best Local Similarity 89.2%; Pred. No. 3.2e-184;
Matches 651; Conservative 36; Mismatches 42; Indels 1; Gaps 1;

QY 1 MLFNSVLRLPOLGVLNRGWSOYPIQSLTLTGQCSCNDEHTSYGETGVVPVPGCTFSSA 60
DB 1 MLFNSVLRLPOLGVLNRGWSOYPIQSLTLTGQCSCNDEHTSYGETGVVPVPGCTFCTA 60
QY 61 PNMEHLAVANEEGFVRLYNTESQSFRRKCFKEMAHNAVDFDLAWVPGLKLVTAAGDQ 120
DB 61 PSMEHLAVANEEGFVRLYNTESQSFRRKCFKEMAHNAVDFDLAWVPGLKLVTAAGDQ 120
QY 121 TAKFWDVAVAGELIGTCKGHQCSLKSVAFSKPKAFVCTCGRDGNIMWTRCNCKKDGFR 180
DB 121 TAKFWDVAVAGELIGTCKGHQCSLKSVAFSKPKAFVCTCGRDGNIMWTRCNCKKDGFR 180
QY 181 QVNOISGAHNTSDKOTPSKPKKQNSKGLAPSDVDFQOSVTVVLFDENTLVSAGAVDGI 240
DB 181 QVNOISGAHNTADKOTPSKPKKQNSKGLAPSDVDFQOSVTVVLFDENTLVSAGAVDGI 240
QY 241 KWDLRKNYTAHQEPIASKFLYPGSSTRKLGYSLLILDSTGSLTFANCTDDNIYFMN 300
DB 241 KWDLRKNYTAHQEPIASKFLYPGSSTRKLGYSLLILDSTGSLTFANCTDDNIYFMN 300
QY 301 TGLKTSFPAIFNGHQNSTFYVKSLSLSPDDQFLVSGSSDEAAIYKWTSTPWPQPTVLLGHS 360
DB 301 TGLKTSFPAIFNGHQNSTFYVKSLSLSPDDQFLVSGSSDEAAIYKWTSTPWPQPTVLLGHS 360
QY 361 QEVTSVCWCPSPDFTKIATCDDNTLKIWLNRGLBEKPGDGLKSTVGWASOKKESRPGL 420
DB 361 QEVTSVCWCPSPDFTKIATCDDNTLKIWLNRGLBEKPGDGLKSTVGWASOKKESRPGL 420
QY 421 VVTSSQSTPAKAPRVKCNPSNPSAACAAPSCAGDPLPSPNTPTFSIKTSPAKARSP 480
DB 421 VVTSSQSTPAKAPRVKCNPSNPSAACAAPSCAGDPLPSPNTPTFSIKTSPAKARSP 480
QY 481 NRRGVSVPSPKPPSPFMSIRNWTPTSSSPPIITPPASETKIMSPRKALIPVSKSSQ 540
DB 481 NRRGVSVPSPKPPSPFMSIRNWTPTSSSPPIITPPASETKIMSPRKALIPVSKSSQ 540
QY 541 AEACSESRNRVRRDLSSCLSVKQKVCSCNCTVLDGQVENLHLDLCLAGNQEDLSK 600
DB 541 AEACSESRNRVRRDLSSCLSVKQKVCSCNCTVLDGQVENLHLDLCLAGNQEDLSK 600
QY 601 DSIQPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLRPLRPGEGSEMVKGKENSSENK 660
DB 601 DSIQPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLRPLRPGEGSEMVKGKENSSENK 660
QY 661 NWLLAAKRAKKAENSPRSPSSQTPNSRRSQSGTKLSPVITPSSMRKICTYFHRKSQED 720
DB 661 NWLLAAKRAKKAENSPRSPSSQTPNSRRSQSGTKLSPVITPSSMRKICTYFHRKSQED 720
QY 721 FCPSEHSTEL 730
DB 721 FCPSEHSTEL 729

RESULT 7
Q80WY1 MOUSE
ID Q80WY1_MOUSE PRELIMINARY; PRT; 732 AA.
AC Q80WY1;
DT 01-JUN-2003, integrated into UniProtKB/TrEMBL.
DT 01-JUN-2003, sequence version 1.
DT 07-FEB-2006, entry version 21.
DE Meth A retinoic-acid regulated nuclear matrix-associated protein.
GN Name=2810047L02Rik; Synonyms=Dtl, Meth A ramp;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
```

OC Muroidea; Muridae; Murinae; Mus.
RN NCBI_TaxID=10090;
RP [1]
RC STRAIN=BALB/c;
RA Skipper J., Shimizu K., Nakayama E.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; AB095736; BAC76405.1; -; mRNA.
DR Ensembl; ENSMUSG00000037474; Mus musculus.
DR MGI; MGI:1924093; 2810047L02Rik.
DR MGI; MGI:1924093; Dcl.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 5.
DR PRINTS; PR00320; GPROTEINBRPT.
DR ProDom; PD000018; WD40; 1.
DR SMART; SM00320; WD40; 5.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_2.
DR PROSITE; PS00682; WD_REPEATS_2; 5.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 732 AA; 79131 MW; 585E6433149F42D0 CRC64;

Query Match
Best Local Similarity 89.0%; Score 3330.5; DB 2; Length 732;
Matches 624; Conservative 36; Mismatches 40; Indels 1; Gaps 1;

QY 1 MLFNSVLROPOLGVLANGHSSQYPLQSLITGTCQSGNDHSTYGETGVPPFGCTFSSA 60
DB 1 MLFNSVLROPOLGVLANGHSSQYPLQSLITGTCQSGNDHSTYGETGVPPFGCTFCTA 60
QY 61 PNWEHLAVANBEGFVRLNNTESQSPKRCFKFEMAHNAWVDFLAWVPELKLVTAAQDQ 120
DB 61 PNWEHLAVANBEGFVRLNNTESQSKTKCFKEMAHNAWVDFLAWVPELKLVTAAQDQ 120
QY 121 TAKFMDVRAKAGELIGTCGHQCSLKSVAFKFAVCTCGRDNIMVMDTRCNKKGDFYR 180
DB 121 TAKFMDVRAKAGELIGTCGHQCSLKSVAFKFAVCTCGRDNIMVMDTRCNKKGDFYR 180
QY 181 QVNOISGAHNTADKQTPSPKPKKQNSKGLAPSVDFQSVTVVLFDQENTLVSAQAVDGI 240
DB 181 QVNOISGAHNTADKQTPSPKPKKQNSKGLAPSVDFQSVTVVLFDQENTLVSAQAVDGI 240
QY 241 KVDLRKNTAYRQEPYASKSFLYPGSSSTRKLGYSLLDSTGSLFANCTDDNIYMFNM 300
DB 241 KVDLRKNTAYRQEPYASKSFLYPGSTRKLGYSLLDSTGSLFANCTDDNIYMFNT 300
QY 301 TGLKTSFPAIFNGHQNSTFYVKSLSLSPDQDFLVSGSSDEAAIYKWTSTPWPPTVLLGHS 360
DB 301 TGLKTSFPAIFNGHQNSTFYVKSLSLSPDQDFLVSGSSDEAAIYKWSMPHPTVLLGHS 360
QY 361 QEVTSVCWCPSDFTKIATCSDDNTLKIWLRLNGLBEPKPGDKLSTVGVASQKKKSRLPGL 420
DB 361 QEVTSVCWCPSDFTKIATCSDDNTLKIWLRLNGLBEPKPGDKLSTVGVASQKKKSRLPGL 420
QY 421 VVTVSQSTPAKAPRVKCNPSNSSPSAACAPSCAGDLPLPNTPTFTSITKSPAKARSP 480
DB 421 VVTVSQSTPAKAPRVKCNPSNSSPSAACAPSCAGDLPLPNTPTFTSITKSPAKARSP 480
QY 481 NRRGSSVSSVSPKPPSKFNSIRNWTPTSSPPPTTTPASSETKIMSPKALIPVSKQSSQ 540
DB 481 NRRGSSVSSVSPKPPSKFNSIRNWTPTSSPPPTTTPASSETKIMSPKALIPVSKQSSQ 540
QY 541 ABACSSSRNRVRRDLSSCLSVKQKCVKSCNCTVLDQOVENHLDLCLAGNQEDLSK 600
DB 541 ABACSSSRNRVRRDLSSCLSVKQKCVKSCNCTVLDQOVENHLDLCLAGNQEDLSK 600
QY 540 ADACSSSRNRVRRDLSSCLSVKQKCVKSCNCTVLDQOVENHLDLCLAGNQEDLSQ 599
DB 540 ADACSSSRNRVRRDLSSCLSVKQKCVKSCNCTVLDQOVENHLDLCLAGNQEDLSQ 599
QY 601 DSLGPTKSSKTEGAGTSISEPPSPVSPYASGCGPLPLRPGEGSEVMVGKENSSENK 660
DB 601 DSLGPTKSSKTEGAGTSISEPPSPVSPYASGCGPLPLRPGEGSEVMVGKENSSENK 660

Db 600 DSEGPTKSSKTEGAGTSISEPPSPVSPYASGCGPLPLRPGEGSEVMVGKENSSENK 659
QY 661 NWLLAAAKRKAENPSPRSPSSQTPNSRQSGKTLPSPTVI 701
DB 660 NWLLAAAKRKAENPSPRSPSSQTPNSRQSGKTLPSPTVI 700

RESULT 8
Q3TTE9 MOUSE
ID Q3TTE9_MOUSE PRELIMINARY; PRT; 597 AA.
AC Q3TTE9;
DT 11-OCT-2005, integrated into UniProtKB/TrEMBL.
DT 11-OCT-2005, sequence version 1.
DT 07-FEB-2006, entry version 6.
DE Adult male testis cDNA, RIKEN full-length enriched library,
DE clone:4930546N03 product:L2DTL protein homolog.
GN Name=Dtl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidea; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=92279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RA "High-efficiency full-length cDNA cloning."
RL Methods Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX PubMed=1641072; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilming L.G., Aidinis V., Allen J.E.,
RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
RA Bansal K.P., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple J.P., De Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
RA Guscinich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Humnietek L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.P., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsumura H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA Mottagui-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavese G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C., Sheng Y.,
RA Schonbach C., Sekiguchi K., Sempile C.A., Seno S., Sessa L., Shiba Y.,
RA Shibata Y., Shimada E., Sugitani K., Sultana R., Takenaka Y., Taki K.,
RA Sperling S., Stupka E., Sugitani K., Sultana R., Takenaka Y., Taki K.,
RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlestedt C., Weiss J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
RA Nishio T., Okada M., Plessy C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watanabe A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome."
RL Science 309:1559-1563(2005).
RN [3]

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis; DOI=10.1101/1101/gr.152600;
RX MEDLINE=20330913; PubMed=11076861; DOI=10.1101/gr.152600;
RG RIKEN Genome Exploration Research Group, and Genome Science Group
RT (Genome Network Core team) and the FANTOM Consortium;
RL "Antisense Transcription in the Mammalian Transcriptome.";
Science 309:1564-1566(2005).
[4]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Dalla E., Bradt D., Brusis V., Ciothia C., Corbani L.E., Cousins S.,
RA Blake E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Giesi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tonita M.,
RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyrshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
[5]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6J; TISSUE=Testis; DOI=10.1038/35055500;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Giesi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Anon H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyrshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
[6]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6J; TISSUE=Testis; DOI=10.1101/1101/gr.145100;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Suganara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630 (2000).

RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Testis; DOI=10.1101/1101/gr.152600;
RX MEDLINE=20330913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-Format
sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771 (2000).
[8]
RN NUCLEOTIDE SEQUENCE.
RP STRAIN=C57BL/6J; TISSUE=Testis;
RA Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
RA Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,
RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Ninomiya N.,
RA Nishiyori H., Nomura K., Ohno M., Sakazume N., Sano H., Sasaki D.,
RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watahiki A.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
EMBL: AK161401; BAE36376.1; -; mRNA.
MGI: MGI:1924093; Dtl.
InterPro: IPR001680; WD40.
Pfam: PF00400; WD40; 4.
PRINTS: PR00320; GPROTEINBRPT.
ProDom: PD000018; WD40; 1.
SMART: SM00320; WD40; 3.
PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_1.
PROSITE: PS00678; WD_REPEATS_2; 4.
PROSITE: PS0294; WD_REPEATS_REGION; 2.
Repeat: WD repeat.
SQ SEQUENCE 597 AA; 64374 MW; A1B8D574ABEF554D CRC64;
Query Match 73.1%; Score 2835.5; DB 2; Length 597;
Best Local Similarity 89.3%; Pred. No. 1.1e-148;
Matches 534; Conservative 28; Mismatches 35; Indels 1; Gaps 1;
QY 133 IGTCGHQCSLKSVAFAKFAVCTGGRDGHIMWDTFCNKGDFYQVQNSGAHNTS 192
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MGTCTGHQCSLKSVAFAKFAVCTGGRDGHIMWDTFCNKGDFYQVQNSGAHNTA 60
QY 193 DKOTPSKPKKKONSKGLAPSVDFQOSVTVLVFDENTLVLSAGAVDGIIVKWLDRKNYAY 252
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 DKQTSPKPKKKONSKGLAPSVDFQOSVTVLVFDENTLVLSAGAVDGIIVKWLDRKNYAY 120
QY 253 ROEPIASKFLYPGSGSTRKGLYSSLLDSTGTLFANCTDDNIYFMNMTGLKTSPIAFIN 312
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 ROEPIASKFLYPGSGSTRKGLYSSLLDSTGTLFANCTDDNIYFMNMTGLKTSPIAVEN 180
QY 313 GHONSTFYKSSLSPPDDQFLVSGSSDEAAIYIKWSTPWPQPTVLLGHSQEVTVSCVCPSD 372
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 GHONSTFYKSSLSPPDDQFLVSGSSDEAAIYIKWSTPWPQPTVLLGHSQEVTVSCVCPSD 240
QY 373 FTKIATCSDDNTLKIWLNRGLKEKPGDKLSTVGWASQKKESRPGCLVTVTSSQSTPAK 432
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 FTKIATCSDDNTLKIWLNRGLKEKPGDKLSTVGWASQKKESRPGCLVTVTSSQSTPAK 299
QY 433 APRVKNPNSNPPSSAACAPSCAGDLPLPSNTPFTSIKTSAPAKRSPINRSGSVSPK 492
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 300 APRAKSSPISPPSSAACTPSCAGDLPLPSNTPFTSVKTTTPATRSSVSRGSSISSVSPK 359
QY 493 PPSFPMSTRNMTPTPSSPPITPPASTTKIMSPKALIPVQSKSQAEACSESNRVK 552
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 360 PLSSPFMSLRNMTPTPSSPPITPPASTTKIMSPKALIPVQSKSQAEACSESNRVK 419

```
QY 553 RRLDSSCLSVKQKVCVNCVTELDQVENLHLDLCLAGNOEDLSKDSLGPTKSSKIE 612
DB 420 RRLDSSCLSVKQKVCVNCVTELDQVAESLRLDCLLSGTQEVLSQDSBEGPTKSSKTE 479
QY 613 GAGTISSEPPSPISPVASESCGTLPLPLPCGEGSEWGVKENSPPKNKWLAMAAKRKA 672
DB 480 GAGTISSEPPSPISPVASESCGTLPLPLPCGEGSEWGVKENSPPKNKWLAMAAKRKA 539
QY 673 ENPSRPSSTQPNRSRQSGKTLPSVTTTPSSMRKICTYFHRKSQEDFCGPEHSTEL 730
DB 540 ENSRPSRPSSTQPNRSRQSGKTLPSVTTTPSSMRKICTYFHRKTDQDFCSPHSTEL 597

RESULT 9
Q6PAN1_MOUSE PRELIMINARY; PRT; 594 AA.
AC O6PAN1;
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.
DT 05-JUL-2004, sequence version 1.
DE 07-FEB-2006, entry version 19.
DE 2810047L02Rik protein.
GN Name=Dtl;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muroidae; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zerbini B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton L., Soares M.B., Donald M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Murry D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Warra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC
CC -----
DR EMBL; BC060208; AAH60208.1; -; mRNA.
DR Ensembl; ENSMUSG00000037474; Mus musculus.
DR MGI; MGI:1924093; Dtl.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 3.
DR PRINTS; PR00320; GPROTEINRPT.
DR ProDom; PD0000018; WD40; 1.
DR SMART; SM00320; WD40; 3.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
DR PROSITE; PS50082; WD_REPEATS_2; 3.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 594 AA; 64233 MW; 180FD4B69B4A823 CRC64;
```

```
Query Match 69.8%; Score 2705; DB 2; Length 594;
Best Local Similarity 87.1%; Pred. No. 1.8e-141;
Matches 514; Conservative 28; Mismatches 42; Indels 6; Gaps 2;

QY 141 CSLSKVAFSKPEKAVFCCTGGRDGNIMWDTTRCNKKDGFYRVNQISGAHNTSDKOTPSKP 200
DB 11 CSLSOVS-----KSCVLYRGERGNIMWDTTRCNKKDGFYRVNQISGAHNTADAKOTPSKP 65
QY 201 KKKQNSKGLAPSVDFQQSVTVVLFQDENTLVLSAGAVDGIWKWDLRKNKYAYRQPIASK 260
DB 66 KKKQNSKGLAPSVDFQQSVTVVLFQDENTLVLSAGAVDGIWKWDLRKNKYAYRQPIASK 125
QY 261 SFLYPGSSSTRKLGYSLLDSTGSTLPAFCTDDNIMFNMTGLKTSPIVAFNGHONSTFY 320
DB 126 SFLYPGTSTRKLGYSLLDSTGSTLPAFCTDDNIMFNMTGLKTSPIVAFNGHONSTFY 185
QY 321 VKSSLSPPDDQFLVSGSSDEAAIYKVKSTPWOPPTVLLGHSEVTSVVCWCPDFTKIATCS 380
DB 186 VKSSLSPPDDQFLVSGSSDEAAIYKVKSWPWPPTVLLGHSEVTSVVCWCPDFTKIATCS 245
QY 381 DDNTLKIWLNRGLLEKPGDKLSTVGWASQKKESRPGLVTVTSSQSTPAKAPVKCP 440
DB 246 DDNTLKIWLNRGLLEKPGDKLSTVGWASQKKESRPGLVTVTSSQSTPAKAPVKCP 304
QY 441 SNSPSSAACAPSCAGDILPLPSNTPTFSIKTSPAKARSPINRSGSVSSVSPKPPSPFKMS 500
DB 305 SISPSAACTPSCAGDILPLPSNTPTFSIKTSPAKARSPINRSGSVSSVSPKPPSPFKMS 364
QY 501 TRNWVTRTPSSPPITPPASETKJMSPRKALIPVSKSQSAEACSESRNRVRRDSSCL 560
DB 365 LKNWVTRTPSSPPITPPASETKJMSPRKALIPVSKSQSAEACSESRNRVRRDSSCL 424
QY 561 ESKVKQKVCVNCVTELDQVENLHLDLCLAGNOEDLSKDSLGPTKSKIEGAGTISIE 620
DB 425 ESKVKQKVCVNCVTELDQVAESLRLDCLLSGTQEVLSQDSBEGPTKSKIEGAGTISIE 484
QY 621 PPSPISPVASSCGTLPLPLPCGEGSEWGVKENSPPKNKWLAMAAKRKAENSPRSP 680
DB 485 PPSPISPVASEGCGPLPLPLPCGEGSEWGVKENSPPKNKWLAMAAKRKAENSPRSP 544
QY 681 SSQTPNSRQSGKTLPSVTTTPSSMRKICTYFHRKSQEDFCGPEHSTEL 730
DB 545 SSQTPNSRQSGKTLPSVTTTPSSMRKICTYFHRKTDQDFCSPHSTEL 594

RESULT 10
Q9NW03_HUMAN PRELIMINARY; PRT; 434 AA.
AC Q9NW03;
DT 01-OCT-2000, integrated into UniProtKB/TrEMBL.
DT 01-OCT-2000, sequence version 1.
DE CDNA FLJ10399 fis, clone NT2RM40J0354, weakly similar to
DE LETHAL(2) DENTICLELESS PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Takekoshi K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y., Ishida S.,
```

RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Inose N., Musashino K., Yuuki H., Ohshima A., Sasaki N., Aotsuka S., Yoshihara Y., Matsunawa H., Ichihara T., Shiohata N., Sano S., Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O., Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H., Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B., Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y., Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii K., Ozaki K., Hirao M., Ohmori Y., Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S., Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M., Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T., Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K., Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R., Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.; RT "Complete sequencing and characterization of 21,243 full-length human cDNAs"; RL Nat. Genet. 36:40-45 (2004). CC ----- CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms> Distributed under the Creative Commons Attribution-NoDerivs License CC ----- CC EMBL; AK001261; BAA91586.1; -; mRNA. DR Ensembl; ENSG00000143476; Homo sapiens. DR InterPro; IPR001680; WD40. DR Pfam; PF00400; WD40; 2. DR PRINTS; PR00320; GPROTEINBRPT. DR ProDom; PD000018; WD40; 1. DR SMART; SM00320; WD40; 2. DR PROSITE; PS00082; WD_REPEATS_2; 2. DR PROSITE; PS0294; WD_REPEATS_REGION; 1. KW Repeat; WD repeat. SQ SEQUENCE 434 AA; 46677 MW; 1B73E2BF1311155D CRC64; Query Match 59.1%; Score 2293; DB 2; Length 434; Best Local Similarity 100.0%; Pred. No. 7.7e-119; Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0; QY 297 MFNMTGLKTSFPAIFNGHQNSTFYVKSLSLSPDDQFLVSGSSDRAAYIKWVSTPWPPTVL 356 DB 1 MFNMTGLKTSFPAIFNGHQNSTFYVKSLSLSPDDQFLVSGSSDRAAYIKWVSTPWPPTVL 60 QY 357 LGHSQEVTSVCWCPSTFTKIATCSDNTLKIWLNRGLEKPGDKLSTVGWASQKKKES 416 DB 61 LGHSQEVTSVCWCPSTFTKIATCSDNTLKIWLNRGLEKPGDKLSTVGWASQKKKES 120 QY 417 RPLGLVTVTSQSTPAKAPRVKCNPSNPSSSAACPSCAGDLPLPSNTPTFSIKTSPAKA 476 DB 121 RPLGLVTVTSQSTPAKAPRVKCNPSNPSSSAACPSCAGDLPLPSNTPTFSIKTSPAKA 180 QY 477 RSPINRRGVSVPKPPSPFKMSIRNWTPTSSGPIPTPASETKIMSPKALIPVSQ 536 DB 181 RSPINRRGVSVPKPPSPFKMSIRNWTPTSSGPIPTPASETKIMSPKALIPVSQ 240 QY 537 KSSQAEACSSRNVRRLDSSCLESVKQKVCVSCNVTDLGQVNLHLDLCLAGNOE 596 DB 241 KSSQAEACSSRNVRRLDSSCLESVKQKVCVSCNVTDLGQVNLHLDLCLAGNOE 300 QY 597 DLSKDSLGPTKSKIEGAGTSISEPPSPISPYASESGTLPPLRPGEGSEMVGKENS 656 DB 301 DLSKDSLGPTKSKIEGAGTSISEPPSPISPYASESGTLPPLRPGEGSEMVGKENS 360 QY 657 PENKNWLLAAAKRKAENPSRPSQTPNSRQSGKTLPSPTVITPSSMRKICTVFHRK 716 DB 361 PENKNWLLAAAKRKAENPSRPSQTPNSRQSGKTLPSPTVITPSSMRKICTVFHRK 420 QY 717 SQEDFCGPEHSTEL 730 DB 421 SQEDFCGPEHSTEL 434

RESULT 11
05ZJW8_CHICK PRELIMINARY; PRT; 720 AA.
AC Q5ZJW8; DT 23-NOV-2004, integrated into UniProtKB/TrEMBL.
DT 23-NOV-2004, sequence version 1.
DT 07-FEB-2006, entry version 10.
DE Hypothetical protein.
GN ORFNames=RCJMB04.15a2;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]_TaxID=9031;
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CB; TISSUE=Bursa;
RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezubov Y., Zaim J.,
RA Fiedler P., Kutter S., Blagodatski A., Kostovska D., Koter M.,
RA Plachy J., Carninci P., Hayashizaki Y., Buerstedde J.M.;
RT "Full-length cDNAs from chicken bursal lymphocytes to facilitate
RT Genefunction analysis.";
RL Genome Biol. 6:R6-R6(2005). CC ----- CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms> Distributed under the Creative Commons Attribution-NoDerivs License CC ----- CC EMBL; AJ270316; CAG31975.1; -; mRNA. DR InterPro; IPR001680; WD40. DR Pfam; PF00400; WD40; 5. DR PRINTS; PR00320; GPROTEINBRPT. DR ProDom; PD000018; WD40; 2. DR SMART; SM00320; WD40; 5. DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_2. DR PROSITE; PS00082; WD_REPEATS_2; 4. DR PROSITE; PS0294; WD_REPEATS_REGION; 1. KW Hypothetical protein; Repeat; WD repeat. SQ SEQUENCE 720 AA; 78633 MW; 1B9101FCA671EB4E CRC64; Query Match 56.9%; Score 2206.5; DB 2; Length 720; Best Local Similarity 61.9%; Pred. No. 8.8e-114; Matches 447; Conservative 73; Mismatches 161; Indels 41; Gaps 12; QY 20 SSQYPLQSLITCVCQSGNDEHTSYGETGVPPFCGTSSAPNMEHVLAVANEEGVRLY 79 DB 16 SSSLPLQLHLDRYCRSDEDDHLSYGEIGMPVFPFGCSFAAPNFEHLVAVANEEGVRLY 75 QY 80 NTESOSFRKKCFKEWMAHNAVFDLAWPGLKLVTAAGDQAKFWDVKAGELIGTCCKGH 139 DB 76 DTEAQTYYKLISKEWQAHNAVFDLAWPGEHRIVTASGDQAKWWDVZAGELLGICKGH 135 QY 140 QCSLKSVAFSKFEKAVFTCTGDRGNIMVMDTRCNCKKDGFYRQVNOISGAHNTSDKOTPSK 199 DB 136 QCSLKSVAFSKFEKAVFTCTGDRGNIMVMDTRCNCKKDGFYRQVNOISGAHNVVDROTPSK 195 QY 200 -PKKQNSKGLAPSVDFQOSVTVVLFDENTLVAGAVDGIKVDLKNRYTAYRQEPKA 258 DB 196 LRKXKQNLRLGLAPLVDFOQSVTVLLQDHTLISAGAVDGVIKVDLKNRYAAYRQDVP 255 QY 259 SKSFLYPGSGSTKGLGVSSLLDSTGLFANCTDDNIYMFNMTGLKTSVPALFNHQNST 318 DB 256 SKSFFYPGTSYTRKGLGVSSLLDSTGLFANCTDDSIYMFNMTGLKTSFPVAVFSGHQNST 315 QY 319 FYVKSLSLSPDDQFLVSGSSDEAAIYIKWVSTPWPPTVLVLGHSQEVTSVCWCPSTFTKIAT 378 DB 316 FYIKSSISLSPDDQFLVSGSSDCNAYIKWVSEPSLPRIILVGHSEVTSIACPSDFTKIAT 375 QY 379 CSDNTLKIWLNR -GLEKPGDKLSTVGWASQKKESRPGLVTVTSQSSTPAKAPRVK 437 DB 376 CSDNTVIRIWLQHYPEEKSVSNKAKLVGWVTKKPEEQRQAGRSASQSTPAKAFSVG 435 QY 438 CNPSNSSPSSAACAPSCAGDLPLPSNTPTFSIKTSPAKARSPINRRGVSVPKPPSPF 497

Qy	121	TAKFDWDVRAGELIGTCKCHQCSLSKVAESKFAVCTGGRDGNIMWMDTRCNCKDKGFYR	180
Db	121	TAKLMDWVWAGEIGSCRHQCSLSKVSFSKPERAVFTSGRDGNIMWMDTRCSCKDKGFYR	180
Qy	181	QVNOISGAHNTSDKTPSKPKKKONS--KGLAPSVDFQOSVTTVVLFDODENTLVSAGAVDGI	239
Db	181	QVNOITGAUNAIDKQTPSKMKRKRPSIRGLAPSVDSQOSVTTVVIFODEYTIISAGAVDGV	240
Qy	240	IKWDLRKNYTAAYROEPIASKSFLYPGSSSTRKGLGYSSLIILDSTGSTLFPANCTDDNIYMFN	299
Db	241	VKIWDLRKNYSAYRQDPVPVKKFPYPGNGSTRKGLGYSSLVLDPTGTNLFASCTDDNYMFN	300
Qy	300	MTGLKTSPPVAFENGHONSTFFVYKSSLPDDQDFLYSGSSDRAAYIKWYSTPQWPTVLLGH	359
Db	301	ATGLKTDVPVSIIFRGHONSTFFIKASVSPDQGFLLSGSSDHSAYIQWVSDPMAAPVTLMGH	360
Qy	360	SOEVTSCVCPSEDFTKIATCSDDNTLKIWLNRGLERKPGDKLSTVGWASQKKESRPG	419
Db	361	COEVTSCVQSDFTKIATCSDDNTVVRWLKRCDESSDSDKRDSCVWACKKFFPS--	418
Qy	420	LVTVTSSTTPAKAPVRVRCNPSNSPASAACAPSCAGDLPLPNSNTPFTSIKTSAPAKARSP	479
Db	419	--SMAANLCTPGKPSVWSSSLTSSPTPASCAPSNCTGDLPMPSSTPISALLPDP-KLQTP	475
Qy	480	--INRGSVSSVSPKPPSPFQMSIRNWWTRTPSSSPPIITPPASSTKIMSPRKALIPVSQK	537
Db	476	KRINNGG--LGVSPKQMSSSKMSIKDWVTRTPKSS-----TRTDKTPSPRKAFTPEQY	528
Qy	538	SSQAEACSE--SRNRVRRRLDSSCLESVKQKVCNCTVELDGOVENLHLDLCCLAGNQ	595
Db	529	PSVSSARVOLPYEKRAKRLRTSS--EYAEHVCPDNCVNCRELEPGLKAKLDVCFI--DK	585
Qy	596	EDLSKDSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLPLRPCGE-GSEMVGKEN	654
Db	586	ERDSDDKRLRLDLSKGFQDELS--FSPSTSLHWNATENL-LQLGPLSELKSVLDDKEN	642
Qy	655	SSPENKNWLLAAWAKRKA--ENPSRPSGSSOTPNRRSQSKTLP-SPVTI--TPSSMRKI	709
Db	643	SSPE-KNWLSALGHKFFKSSPQNKASGSPSRTSTTKKQOPRNAPNSPVSPYTPPGSMRKI	701
Qy	710	CTYFHRKSQ 718	
Db	702	CTYFFKCKSE 710	

RESULT 14

Q6GPU3_XENIA	PRELIMINARY; PRT; 711 AA.
ID	Q6GPU3_XENIA PRELIMINARY; PRT; 711 AA.
AC	Q6GPU3;
DT	19-JUL-2004, integrated into UniProtKB/TrEMBL.
DT	19-JUL-2004, sequence version 1.
DT	07-FEB-2006, entry version 16.
DE	MC82606 protein.
GN	Name=MC82606;
OS	Xenopus laevis (African clawed frog).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC	Xenopodinae; Xenopus; Xenopus.
OX	NCBI_Taxid=8355;
RP	[1]
RN	NUCLEOTIDE SEQUENCE.
RC	TISSUE=Ovary;
RC	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.
RA	Klauser R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA	Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA	Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Plange C.,
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Liu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skaleka U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2].
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Ovary;
RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/advy.10174;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative";
RL Dev. Dyn. 225:384-391 (2002).
RN [3].
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Ovary;
RA Klein S., Gerhard D.S.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; BC073015; AAH73015.1; -; mRNA.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 5.
DR PRINTS; PR00320; GPROTEINBRPT.
DR ProDom; PD000018; WD40; 2.
DR SMART; SM00320; WD40; 4.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_2.
DR PROSITE; PS00682; WD_REPEATS_2; 5.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 711 AA; 78124 MW; D68C75007C521F33 CRC64;

Query Match 52.5%; Score 2034.5; DB 2; Length 711;
Best Local Similarity 56.8%; Pred. No. 2.8e-104;
Matches 423; Conservative 86; Mismatches 175; Indels 61; Gaps 18;

QY 1 MLFNSVLROPLQGLVLRNGSSQYPLQSLTGYOCNGNDHTSYGETGVVPFGCTFSSA 60
DB 1 MLFRSVNKPFRFCGCHNGAPFTSPLOSLQCYQVQKDBSHISYGLDGRAVPVPGCAFTV 60
QY 61 PNMEHVLAVANEFGFVRLYNTESQSFRKCKFEKMAHNAVDFLAWVPGELKLVTAAGDQ 120
DB 61 PGMSHVLAVANEFGVRLYDTECGDMQRLVVKFEFMAHTNAVDFIAVWPGHEKLVTAAGDQ 120
QY 121 TAKFMDVKAGELIGTCKGHQCSLSVAFSKFEKAVFCTGGRDGNIMVWTRCNKKDGFYR 180
DB 121 TAKLMDVKAGELIGECKGHQCSLSVAFSKFERAVFSTGGRDGNIMVWTRCNKKDGFYR 180
QY 181 QVNQISGAHNTSDKQTPSKPKKQNS - KGLAPSVDPFQQSVTVVLFODENTLVSAGAVGI 239
DB 181 QVNQITGAHNAIDKQTPSKVKRKPISIRGLAPSVDSQQSVTVVIFQDEYTVISAGAVDV 240
QY 240 IKVWDLRKNYAVRQEPPIASKSFLYPSSTRKGLYSLLIDSTGSLFANCTDDNLYMEN 299
DB 241 VKIWDLRKNYSTYRQDPVPVKLPFGNSTRKGLYSLLVDPFTGTLNLFASCTDDNLYMEN 300
QY 300 MTGLKTPSPAIFNGHONSTFYVKSSLSPPDQFLVSGSSDEAAAYIKVSTFWPQPTVLLGH 359
DB 301 ATGLKTDPSVFRGHQNSIFYIKASVSDQFLSGSSDHSAYIMQVSDPLAAPINLIGH 360
QY 360 SQEVTSVCWCPDFTKIATCSDNTYKIWLNRGLBEKPGDGKLSVTGWAASOKKESRPG 419
DB 361 CQEVTSVAVWCSDFTKIATCSDNTYKIWLNRGLBEKPGDGKLSVTGWAASOKKESRPG 418
QY 420 LVTVTSSQSTPAKPRVKCNPSNSPSFSSAACAPSCAGDLPLPNTPTFS - IKTSPAKRS 478

DB 419 --SMAAILCTGKPKSMIPSSSLMSSPTPATCAPSNTGDLPLPSSSTPLSALLPTSKLQTPK 476
QY 479 PINRRGVSVPKPPSPFKMSIRNVWTRTPSSSPPTTPPASETKIMSPKALIPVSQKS 538
DB 477 RINNGG--LGASPKQMSSSKISIKDWVTRTPKSS-----KGTDSKTPSPKRAFTVEQYP 529
QY 539 --SQAEACSESRNVKRLDSSCLSEVKQKCVKSCNCVTELDGQVENLHLDLC----- 589
DB 530 IVSNARVOLPYEKRAKRRLTSS--EDVEHVCLDNC-CVMELEPRLLKSKLDLCSLNERDC 587
QY 590 ----CLAGNOBDSLKD----SLGPTKSSKIEGAGTISISEPPSPISPYASESCOTLPLPLRP 642
DB 588 RDDKCL--RLSDLSKEFQOELSPSPSTSLHWNAT--DNPPT-----LSP 627
QY 643 CGE--GSEVMGKENSPEKNWLLAWAAKRAENSPSPSPSQTNSRRQSK-----TLP 696
DB 628 LSEMKSDFDVKENSPE-KNWLALGHKFKSDNSPQFKSSSSPQKSSSSPSRSTSKKHQPNAP 686
QY 697 -SPVTI--TPSSMRKICTYFHRKSQ 718
DB 687 NSPVSVPTTTPGMRKICTYFFRKSE 711

RESULT 15
QSRH15 BRARE
ID QSRH15_BRARE PRELIMINARY; PRT; 647 AA.
AC QSRH15;
DT 21-DEC-2004, integrated into UniProtKB/TrEMBL.
DT 21-DEC-2004, sequence version 1.
DT 07-FEB-2006, entry version 9.
DE Denticleless homolog (Drosophila).
GN Name=dtl; ORFNames=DKFY-18C8.4-001;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Beasley H.;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; BX510324; CAI20732.1; JOINED; Genomic_DNA.
DR EMBL; BX511163; CAI20732.1; -; Genomic_DNA.
DR EMBL; BX510324; CAI20792.1; -; Genomic_DNA.
DR EMBL; BX511163; CAI20792.1; JOINED; Genomic_DNA.
DR Ensembl; ENSDARG00000023002; Danio rerio.
DR ZFIN; ZDB-GENE-020419-34; dtl.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 5.
DR PRINTS; PR00320; GPROTEINBRPT.
DR ProDom; PD000018; WD40; 2.
DR SMART; SM00320; WD40; 5.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_2.
DR PROSITE; PS00682; WD_REPEATS_2; 5.
DR PROSITE; PS0294; WD_REPEATS_REGION; 1.
KW Repeat; WD repeat.
SQ SEQUENCE 647 AA; 70769 MW; 33A45A01B42621DC CRC64;

Query Match 42.1%; Score 1630.5; DB 2; Length 647;
Best Local Similarity 48.5%; Pred. No. 5.7e-82;
Matches 355; Conservative 85; Mismatches 163; Indels 129; Gaps 20;

QY 16 RKGWSSQYPLQSLTGYOCNGNDHTSYGETGVVPFGCTFSSAPNMEHVLAVANEFG 75
DB 17 RKGQRPLFPLSLDGYECARDEHISIGASAAVPPFGCTFSSAHGQONCLAVANEFG 76
QY 76 VRLYNT-ESQSPRKKCFKEMAHNAVDFLAWVPGELKLVTAAGDQTKAFMDVKAGELIG 134

```

Db      77 VTIFNTGEKQS---SVLKEWQAHDNAVFDAIAWVPGTNCVLWTASGDQOTARLWDVITGDLLG 133
Qy      135 TCKGHQCSLKSVAFSKFEXVCTGGRDGNIMVWDTKCNKKGDFYOVNOISGAHNTSDK 194
Db      134 TFKGHQCSLKSVAFYKQEKAVFSTGGRDGNIMIWDTRCSKKGDFYQVQKQISGAHWKPER 193
Qy      195 QTPSKPKKKQNSKGLAPSVDFQOSVTVVLFDQDENTLVLSAGAVDGIKVVWDLRKNYTYARQ 254
Db      194 FTPQTKR---RGMAPPVDSQGVTVLFCDETKLISSGAVDGIKVVWDLRKNYTYAHQ 249
Qy      255 EPIASKFLYPGSTRKGLYSLILDSTGSTFLPANCDDNIYMFNMTGLKTSFVALFNCH 314
Db      250 NPLPLQAYPYPGSCTRKGLYSLGLDYGSRFLFSNCTDDNIYMFNISGLKTTFPAVAFSGH 309
Qy      315 QNSTFVVKSLSDDDQFLVSGSSDEAAYIWKVSTPMQPTVLLGHSEVTSVCWCPSDFT 374
Db      310 SNSSFYKSVSPDDQFLASGSSDHNVIWKISDPKQAPMMLQGHSEVTSVAMCPTDFT 369
Qy      375 KIATCSDDNTLKIRLNRGLEEKPGGDKL-----STVGWASQKKESR--PGLVTVTSQ 427
Db      370 KIASCSDNTVRIWRLNR---KPEGENSTIQDGNLVGWTIRKQVSPNRTPG--HHSPE 423
Qy      428 STPAKAPRVKCNPSNPSAACAACAPSCAGDLPLPSNTPFTSFKTSAPAKARSPINRGVS 487
Db      424 LTPSKNPGSVRSVLASPOPATCAPTCAA-LPLPSNTSS---APPAKL-----467
Qy      488 SVSPKPPSPFKMSIRNWTPTSSSPITPPASETKIMSPKALIPVSKSSQAEACSES 547
Db      468 -TSPKMP-----SLOQWISR--SSKSPV-----RKALTPVLOGLS-----F 501
Qy      548 RNRVRRLLDSSCLESVKQKVCNCVTELDQOVENLHLDLCLAGNQEDLSKDSLGPTK 607
Db      502 EHRVXRLETG-----DSASSGLG---520
Qy      608 SSKIEGAG---TSISEPPPIPIYASESCGTLPLRPGEGSEMVGKNSPENKNWLL 664
Db      521 -EIDGVSELYPNVKRSRVSSTLKKEDSFGLESEKRLGSDGAESAEGKNSPRRTDWLS 579
Qy      665 ANNAKKAENPSPRSPSSQTPNRRSGKTLPSPTITP-----SSMR 707
Db      580 VISQKFKG-SAQPKSPSS---GSSQODTRTLESAAVSPRPMKVFSPTNKKASPSKPMK 635
Qy      708 KICTYFHRKSQE 719
Db      636 KISSYFWKRTQD 647

```

Search completed: November 8, 2006, 18:13:51
 Job time : 315 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2006, 18:15:55 ; Search time 44 Seconds
(without alignments)
1408.400 Million cell updates/sec

Title: US-10-726-160-2

Perfect score: 3877

Sequence: 1 MLFNSVLRQPQLGVLRNGWS.....TYFHRKSQEDFCGPEHSTEL 730

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 326800 seqs, 84889852 residues

Total number of hits satisfying chosen parameters: 326800

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA New:*

- 1: /EMC_Celerra_SID33/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 2: /EMC_Celerra_SID33/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 3: /EMC_Celerra_SID33/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 4: /EMC_Celerra_SID33/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 5: /EMC_Celerra_SID33/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 6: /EMC_Celerra_SID33/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 7: /EMC_Celerra_SID33/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 8: /EMC_Celerra_SID33/ptodata/1/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3503.5	90.4	673	6	US-10-540-898-674
2	2914.5	75.2	784	6	US-10-540-898-671
3	759	19.6	146	6	US-10-540-898-676
4	267	6.9	321	6	US-10-953-349-33327
5	255	6.6	278	6	US-10-953-349-33329
6	255	6.6	279	6	US-10-953-349-33328
7	250	6.4	415	7	US-11-293-697-4324
8	249	6.4	301	6	US-10-449-902-36351
9	241	6.2	343	6	US-10-449-902-56313
10	240	6.2	244	6	US-10-449-902-41812
11	234	6.0	317	7	US-11-056-355B-90488
12	234	6.0	317	7	US-11-056-355B-94244
13	231.5	6.0	319	7	US-11-056-355B-13097
14	231.5	6.0	346	7	US-11-056-355B-13096
15	231.5	6.0	1281	6	US-10-449-902-42332
16	227	5.9	808	6	US-10-449-902-41350
17	219	5.6	324	6	US-10-449-902-36608
18	217.5	5.6	299	7	US-11-056-355B-23167
19	217.5	5.6	299	7	US-11-056-355B-77008
20	217.5	5.6	299	7	US-11-056-355B-79227
21	217.5	5.6	302	7	US-11-056-355B-84740
22	217.5	5.6	302	7	US-11-056-355B-77007
23	217.5	5.6	304	7	US-11-056-355B-23166
24	217.5	5.6	304	7	US-11-056-355B-84739
25	217.5	5.6	837	7	US-11-056-355B-108275

ALIGNMENTS

RESULT 1

US-10-540-898-674
; Sequence 674, Application US/10540898
; Publication No. US20060166213A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: CHIR0056-101 (PP023367.0003)
; CURRENT APPLICATION NUMBER: US/10/540,898
; PRIOR FILING DATE: 2005-06-27
; PRIOR APPLICATION NUMBER: US 10/330,773
; PRIOR FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 674
; LENGTH: 673
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-540-898-674

Query Match 90.4%; Score 3503.5; DB 6; Length 673;
Best Local Similarity 91.8%; Pred. No. 1e-185;
Matches 670; Conservative 1; Mismatches 2; Indels 57; Gaps 2;

QY	1	MLFNSVLRQPQLGVLRNGWSQYPIQSLITGTQCSGNDHTSYGTVGPVPPGCTFSSA	60
DB	1	MLFNSVLRQPQLGVLRNGWSQYPIQSLITGTQCSGNDHTSYGTVGPVPPGCTFSSA	60
QY	61	PNNEHVLA VANE EGFVRLYNTESQSFRRKCKPEWMAHNAVFDLAWVPGELKLVTAAGQ	120
DB	61	PNNEHVLA VANE EGFVRLYNTESQSFRRKCKPEWMAHNAVFDLAWVPGELKLVTAAGQ	120
QY	121	TAKFMDVKAGELIGTCKGHQCSLKSVAFSKPKAVFCTGGRDGNIMVWTRCNKXKDGFR	180
DB	121	TAKFMDVKAGELIGTCKGHQCSLKSVAFSKPKAVFCTGGRDGNIMVWTRCNKXKDGFR	180
QY	181	QVNIQISGAHNTSDKQTPSPKPKKQNSKGLAPSVDFQOSVTVVLFQDENTLVAGAVDGI	240
DB	157	-----DFQQSVTVVLFQDENTLVAGAVDGI	183
QY	241	KVWDLRKNYAVRQPIASKSFYPGSSRKLKGLYSSLLDSTGSLFANCTDDNIYFNM	300
DB	184	KVWDLRKNYAVRQPIASKSFYPGSSRKLKGLYSSLLDSTGSLFANCTDDNIYFNM	243
QY	301	TGLKTSFPAIFNGHQNSTFYVKSLSLSPDQFLVSGSSDEAAIWKVSTPWQPTVLLGHS	360
DB	244	TGLKTSFPAIFNGHQNSTFYVKSLSLSPDQFLVSGSSDEAAIWKVSTPWQPTVLLGHS	303

Db 61 CLESVQKCVKNCVTELDQGVENLHLDLCCLAGNQEDLSKDSLGPTKSSKIEGATSI 120
QY 619 SEPPSPISPYASESCGTLPLPLRPC 643
Db 121 SEPPSPISPYASESCGTLPLPLRPC 145

RESULT 4
US-10-953-349-33327
; Sequence 33327, Application US/109533349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33327
; LENGTH: 321
; TYPE: PRP
; ORGANISM: Zea mays subsp. mays
US-10-953-349-33327

Query Match 6.9%; Score 267; DB 6; Length 321;
Best Local Similarity 24.7%; Pred. No. 1.2e-07;
Matches 98; Conservative 61; Mismatches 140; Indels 98; Gaps 15;
QY 58 SSAPNMEHLAVANBEGFVRLYNTESQSFRRKCKFKEMAHWNAVFDLAWVPGLKLVATA 117
Db 11 SASNDKSAIIWEDEDEGLLRHLTL-----GHRKSVMMVWVSPDDCQLLTG 58
QY 118 GDQAKFWDVYKAGELICTKGHCQSLKSVAFSKFAVFCCTGGRDGNIMVWTRCNKXDG 177
Db 59 QBETIRRDVKSCKLHVYKESATGLISCW--FPDGKQLSGLADQNFQCIWDLGKEVDC 116
QY 178 FYRQVNOISGAHNTSKDTPSK-----PKKQNSKGLAPSDVDFQVSVVFLQDENTLVS 232
Db 117 WKQD-----GSTGSDFAVAKDGNLIISMSKQ-----TILLFDRET----- 153
QY 233 AGAVDGLIIVKDLRKNVAVRQPIASKSLYPSSTRKLGYSLLDSTGSLFANCTD 292
Db 154 -----KQRIIE-----GST-----ITSFCLSEDGDFLLVNLVS 183
QY 293 DNIYFMNTGLKTSFPAI--FNHQNSTFYVKSLSLSPDDQ-FLVSGSDEAAIYIKVSTP 349
Db 184 BEIHLWN---IRNDPVRVNRVNGHKTRFVIRSCFGPEQAFIASCSDSQVYIHRATG 240
QY 350 WQPTVLLGHQSVTSVWCPSDFTKIATCSDNTLKIWLNRGLBEKPGDKLSTVGWA 409
Db 241 DLITET-LAGHSGTVNCVSWNPANPHMLASASDDHTVRIW-----GAKK-----S 283
QY 410 SOKKESRPLVTVTSQSSTPAKAPV---KCNPSNS 443
Db 284 SLKRKDVSSNCNSNGSRNGNAHNVFVQKCGNNT 320

RESULT 5
US-10-953-349-33329
; Sequence 33329, Application US/109533349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 33329
; LENGTH: 278
; TYPE: PRP
; ORGANISM: Zea mays subsp. mays
US-10-953-349-33329
Query Match 6.6%; Score 255; DB 6; Length 278;
Best Local Similarity 25.6%; Pred. No. 4.5e-07;
Matches 90; Conservative 51; Mismatches 124; Indels 86; Gaps 14;
QY 104 LAWVPGELKLVTAAGDQTAKEFWDVYKAGELICTKGHCQSLKSVAFSKFAVFCCTGGRDG 163
Db 2 VAWSPDDCQLLTGQEBETIRRDVKSCKLHVYKESATGLISCW--FPDGKQLSGLADQ 59
QY 164 NIMVWTRCNKXDGFYRQVNOISGAHNTSKDTPSK-----PKKQNSKGLAPSDVDFQOS 218
Db 60 NFCIWDLDGKEVDCWKQ-----GSTGSDFAVAKDGNLIISMSKQ----- 101
QY 219 VTVLFDQDENTLVSAGAVDGLIIVKDLRKNVAVRQPIASKSLYPSSTRKLGYSLLI 278
Db 102 -TILLFDRET-----KQRIIE-----GST-----ITSFC 126
QY 279 LDSTGSLFANCTDDNIYFMNTGLKTSFPAI--FNHQNSTFYVKSLSLSPDDQ-FLVSG 335
Db 127 LSEGDGDFLLVNLVSEIHLWN---IRNDPVRVNRVNGHKTRFVIRSCFGPEQAFIASC 183
QY 336 SDEAAIYIKVSTPWPQPTVLLGHQSVTSVWCPSDFTKIATCSDNTLKIWLNRGL 395
Db 184 SEDSQYIWHRATGDLIET-LAGHSGTVNCVSWNPANPHMLASASDDHTVRIW----- 235
QY 396 EKPGDGKLSVGVASOKKESRPLVTVTSQSSTPAKAPV---KCNPSNS 443
Db 236 ---GAKK-----SLKRKDVSSNCNSNGSRNGNAHNVFVQKCGNNT 277

RESULT 6
US-10-953-349-33328
; Sequence 33328, Application US/109533349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; NUMBER OF SEQ ID NOS: 40252
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33328
; LENGTH: 279
; TYPE: PRP
; ORGANISM: Zea mays subsp. mays
US-10-953-349-33328

Query Match 6.6%; Score 255; DB 6; Length 279;
Best Local Similarity 25.6%; Pred. No. 4.5e-07;
Matches 90; Conservative 51; Mismatches 124; Indels 86; Gaps 14;
QY 104 LAWVPGELKLVTAAGDQTAKEFWDVYKAGELICTKGHCQSLKSVAFSKFAVFCCTGGRDG 163
Db 3 VAWSPDDCQLLTGQEBETIRRDVKSCKLHVYKESATGLISCW--FPDGKQLSGLADQ 60
QY 164 NIMVWTRCNKXDGFYRQVNOISGAHNTSKDTPSK-----PKKQNSKGLAPSDVDFQOS 218
Db 61 NFCIWDLDGKEVDCWKQ-----GSTGSDFAVAKDGNLIISMSKQ----- 102
QY 219 VTVLFDQDENTLVSAGAVDGLIIVKDLRKNVAVRQPIASKSLYPSSTRKLGYSLLI 278
Db 103 -TILLFDRET-----KQRIIE-----GST-----ITSFC 127
QY 279 LDSTGSLFANCTDDNIYFMNTGLKTSFPAI--FNHQNSTFYVKSLSLSPDDQ-FLVSG 335
Db 128 LSEGDGDFLLVNLVSEIHLWN---IRNDPVRVNRVNGHKTRFVIRSCFGPEQAFIASC 184

QY 336 SDEAAYIKWSTPWPPTVLGHQSOEVTVCWCSDFDKIATCSDDDNTLKIWRNLRGLE 395
Db 185 SEDSQYIWHRTAGDIET-LAGHSGTNCVGNFANPHMLASASDDHTVRIW----- 236
QY 396 EXPGGDKLSTVGWASOKKESRPLVTVTSOSTPAKAPRV---KCNPSNS 443
Db 237 ---GAKK-----SSLKRKDVSSNCNSGRSNGNAHNVFQKNGNNT 278

RESULT 7

US-11-293-697-4324
; Sequence 4324, Application US/11293697
; Publication No. US20060105376A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/11/293,697
; CURRENT FILING DATE: 2005-12-05
; PRIOR APPLICATION NUMBER: US/10/108,260
; PRIOR FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4324
; LENGTH: 415
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-293-697-4324

Query Match 6.4%; Score 250; DB 7; Length 415;
Best Local Similarity 24.3%; Pred. No. 1.3e-06;
Matches 89; Conservative 58; Mismatches 135; Indels 84; Gaps 14;

QY 35 SGNDHTSGTGVPVP-----PFGCTSSAPNMEHVLAVANEGBGVRLYNTESQSPRKK 89
Db 123 SGEELNTLEGRNVVYAIAPNPYG-----DKIATGSPDKLWSVET-----CK 168
QY 90 CFKEMAHNAVPDLAWPGEIKLVTAAGDQAKFMDVKVAGELIGTCCKHQCSLKVAFS 149
Db 169 CYHTFRGTAETVCLSFNQSLVATGSMDDTAKLWDIQNGBEVTVLRGHSAAEIIISFN 228
QY 150 KPEKAVCTCGRGNIMVMDTRCNKDKGYRQVNOISGAHNTSDKQTPSKPKKQNSKGL 209
Db 229 TSGDRII-TGSEFDHTVVVVDATGRK-----VNLIIG----- 259
QY 210 APSVDQQSVTVVLFODENTLVSAGAVDGIIVKMDLRKNYATYRQPIASKSFLYPGSST 269
Db 260 -----HCAEISSAFNWDCLSLITGSMCKTLWD-----ATNGKCVATLT----- 300
QY 270 RKLGYSSLILDS----TGSTLFANCTDDNIYFMNMTGLKTSPTVAIFNGHQNSTFFVYKSSL 325
Db 301 ---GHDDEILDSCFDVTGKLIATASADGTARIFSAATKRC--IAKLEGHEGE--ISKISF 353
QY 326 SPDDQLVSGSDAAIYIKWSTPWPPTVLGHQSOEVTVCWCSDF--TKIATCSDDN 383
Db 354 NFGQNHLLTGSSDKTARIWDAQT--GQCLQVLEGHTDEIFS---CAPNYKGNIVITGSKDN 409
QY 384 TLKIWR 389
Db 410 TCRWR 415

RESULT 8

US-10-449-902-36351
; Sequence 36351, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF

; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36351
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-36351

Query Match 6.4%; Score 249; DB 6; Length 301;
Best Local Similarity 25.4%; Pred. No. 1.1e-06;
Matches 86; Conservative 51; Mismatches 124; Indels 78; Gaps 13;

QY 66 VLAVA-NEEG-----FVRLYNTESQSFRRKKCFKEMAHNAVPDLAWPGEIKLV 114
Db 24 VLAVFRNDRGNYCLSCGKDRIRLWNPHTGAL----VKPYKSHGREVRDVMSSSDNAKLV 79
QY 115 TAAGDQAKFMDVKAGELIGTCCKHQCSLKVAFSKFKAFTCTGGRGNIMVMDTRCNK 174
Db 80 SCGDRQVYWDVASARVIRKFRGHSEINSEVKNF--NTVVVSAGYDRTVRAFDCR--- 135
QY 175 KDFYRQVNOISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQSQVTVVLFODENTLVSAG 234
Db 136 -----SQSSDPIQTIDTFQDSVMSVNL--TNTETIAG 165
QY 235 AVDGLIKWDLRKNYATYRQPIASKSFLYPGSSTRKLG--SSILLDSTGTLFANCTD 292
Db 166 SVDTGRTIFDRMG-----RETVNLIHGHPVNCISLSDNRNCLLANCLD 208
QY 293 DNIYFMN-MTGLKTSPTVAIFNGHQNSTPYKSSLSPPDQFLVSGSSDAAIYIKWSTPQW 351
Db 209 STVRLDKSTG---ELLOEYKGHICKSPKMDCCLTNDDAFVVGSGSEDGYIFFWELV---D 262
QY 352 PPTV--LLGHQSOEVTVCWCSDFTKIATCSDDNLTKIW 388
Db 263 APVSSPRAHSSVTVTSVYHPTRACML--TSSVDGTIRVW 300

RESULT 9

US-10-449-902-56513
; Sequence 56513, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56513
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-449-902-56513

Query Match 6.2%; Score 241; DB 6; Length 343;
Best Local Similarity 23.4%; Pred. No. 3.4e-06;
Matches 74; Conservative 52; Mismatches 114; Indels 76; Gaps 10;

```
QY 92 KEMW---AHNNAVFDLAWPGLKLVTAAGDQTAKEFDWVKAGBLIGTCCKGHCQSLKSVAF 148
Db 87 KNYVVLRGHKNVLDLQWITDGTQIISAPDKTVRVWDVETGKQVKMAEHSSFYNSCCP 146
QY 149 SPEKAVFCTGGRDGNIMWDTFCNKKDGFYRQVQISGAHNTSDKQTPSKPKKQNSKG 208
Db 147 ARKWPLVVGSGDGTAKLWDLR-----QRG 172
QY 209 LAPSVDFQOSQVTVLVFQDENTLVSAAGVGLIKVWDLRKN-VTAYRQEPISKSLFPLPGS 267
Db 173 AIGTLPDKVQITAVSFSEADKRVFTGGLDNDVWDLRKNNEVEYLLK-----219
QY 268 STRKLGYSLLI-----LDSTGSTLFANCTDDNIYFNN-----TGLKTSPPVAFNGHQN 316
Db 220 -----GHQDMITGQWLSPOGSSYLLTNAMONELKIWDLRPVAPENRIKT-----LTHQH 269
QY 317 S--TFVVKSLSPDDQFLVSGSDSEAAAYTWKSTPWOPPTVLLGHGQEVTSVCWCPSDFT 374
Db 270 NFEKNLLKCSWSPDNKRVTAGSADRMVYIWD-TTSRRILYKLFHNGSVNETAFHPTE-P 327
QY 375 KIATCSDDNTLKIWRL 390
Db 328 VIGSCGSDKQIYLDEL 343
```

```
RESULT 10
US-10-449-902-41812
; Sequence 41812, Application US/10449902
; Publication No. US20060123505A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205Y1-US
; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 41812
; LENGTH: 244
; TYPE: prt
; ORGANISM: Oryza sativa
US-10-449-902-41812

Query Match 6.2%; Score 240; DB 6; Length 244;
Best Local Similarity 29.1%; Pred. No. 2.6e-06;
Matches 65; Conservative 36; Mismatches 86; Indels 36; Gaps 6;

QY 307 PVAIFNGHONSTFYVKSLSPPDQFLVSGSDSEAAAYIKVSTPWQPTVLLGHGQEVTSV 366
Db 5 PIKAYTGSKIESFVKSAISPDGTHILGSSDGNVILWQDQPERGPILLEGEGEATSV 64
QY 367 CWCPSDFTKIATCSDDNTLKIWRNLGRLEKPGGDKLSTVGWASQKKESRPGLVTTSS 426
Db 65 DWCASEVGKIATSDSDSKVRVMTERRV--FPNTSSPTVI-----RKRTAPN-----TGS 113
QY 427 QSTPAKAPRVKCNPSNPSAACAISCAGDLPLPNTP-----TFSIKTSPAK-----475
Db 114 RSASHEL-----ATTSRDYGVAACTSDAGELTGSRLPQVRLEFGTPEAKKAFRL 167
QY 476 -----AFSPINRGVSVSVPKPPSPKMSIRNVTFTPS 510
Db 168 FOEDSLDIRKSPQAQMNPSVSLSPPHSLKRRITRDIYFASSSS 210
```

RESULT 11
US-11-056-355B-90488

```
; Sequence 90488, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 90488
; LENGTH: 317
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; LOCATION: (1)..(317)
; OTHER INFORMATION: Ceres Seq. ID no. 12724103
US-11-056-355B-90488

Query Match 6.0%; Score 234; DB 7; Length 317;
Best Local Similarity 22.7%; Pred. No. 7.5e-06;
Matches 75; Conservative 66; Mismatches 126; Indels 64; Gaps 11;

QY 66 VLAVANEEGFRLY--NTESQSFRRKCKPEKMAHNAVFDLAWVPGELKLVTAAGDOTAK 123
Db 38 LLASASADKTIRTYTINTINDPIAP-VQEFYTGHEGHSIDVAFSSDARFIVASDDKTLK 96
QY 124 FWDVKAAGELIGTCCKGHCQSLKSVAFSKPEKAVFTCTGGRDGNIMWMDTRCNKDKGFYRQV 183
Db 97 LMDVETGSLIKTLIGHTNYAFCVFNFP-QSNMIVSGSPDETVRINDVTTGK-----CU 148
QY 184 QISGAHNTSDKQTPSKPKKQNSKGLAPSDVPQQSVTVLVFQDENTLVSAAGVAGIIVKW 243
Db 149 KVLPAHS-----DPTAVDFNRDGLSVSSSVYDGLCRIW 182
QY 244 D-----LRKNYATYRQEPISKSLFPLPGSSTRKLGYSLLDSTGTLFANCTDDNIYF 298
Db 183 DSGTGHCVKTLIDDENPPVSVRF-----SPNGKFIILVGLTLDNTLRLW 225
QY 299 NMTGLKTSPPVAFNGHONSTFYVKSLS-PDQFLVSGSDSEAAAYIKVSTPWQPTVLL 357
Db 226 NISSAKF--LKTYTGHVNAQYCISSAFSVTNGKRVISGSDCNVHWELNSK-KLLQKLE 282
QY 358 GHSQEVTSVCWCPSPDFTKIATCSDDNTLKIW 388
Db 283 GHTETVMNVACHPTE-NLIASGSLDKTVRIW 312

RESULT 12
US-11-056-355B-94244
; Sequence 94244, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 94244
; LENGTH: 317
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
```

```
; LOCATION: (1)..(317)
; OTHER INFORMATION: Ceres Seq. ID no. 12724103
US-11-056-355B-94244

Query Match
Best Local Similarity 6.0%; Score 234; DB 7; Length 317;
Matches 75; Conservative 66; Mismatches 126; Indels 64; Gaps 11;

QY 66 VLAVANEGBFVRLY--NTSQSFRKCKFKEMAHNAVFDLAWVPGELKLVTAAGDQAK 123
DB 38 LLASASADKTIIRTYTINTINDPIAEP-VQEFTHENGISDVAFSSDARPIVSADDTLK 96
QY 124 FWDVAGELIGCKGHQCSLSKVAFSKFEKAVFCTGGRGNIMWDTNCKKDGFRQVN 183
DB 97 LMDVETGSLIKTLIGHTNYAFVNFNP-QSNMIVSGSFDETIRIWDVTGK-----CL 148
QY 184 QISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQOSVTVVLFDENTLVSAGAVDGIKVM 243
DB 149 KVLPAHS-----DPTAVDFNRDGLSVSSVSDGLCRIM 182
QY 244 D-----LRKNYATAYROEPIASKSFLYPGSTRKLGYSLLDSTGTLFANCTDDNIYMF 298
DB 183 DSGTGHCVKTLIDDENPPVSFVR-----SPNGKFIILVGTLDNTLRIM 225
QY 299 NMTGLKTSVPAFNCHQNSTFYVKSSLS-PDQFLVSGSSDEAAIYIKVSTPWQPTVLL 357
DB 226 NTSSAKF--LKTYTGHVNAQYCISSAFSVTNGKRIVSGSEDNCVHMWELNSK-KLLQKLE 282
QY 358 GHSQEVTSVCMWPSDFTKIATCDDNTLKIW 388
DB 283 GHTETVMNVACHPTE-NLIASGLDKTVRIW 312

RESULT 13
US-11-056-355B-13097
; Sequence 13097, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: 2005-02-14
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 13097
; LENGTH: 319
; TYPE: prt
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(319)
; OTHER INFORMATION: Ceres Seq. ID no. 12344627
US-11-056-355B-13097

Query Match
Best Local Similarity 6.0%; Score 231.5; DB 7; Length 319;
Matches 84; Conservative 63; Mismatches 142; Indels 93; Gaps 14;

QY 20 SSQYPLQSLITGYQCSGNDEHTSYGETGVPPVPPGCTFSSAPNMEHVLAVANEGBFVRLY 79
DB 12 SPGYVLRSTLEGR-----RAVSTVKFSPDG-----RLLASASADKLLRW 52
QY 80 NTESQSRKCKFKEMAHNAVFDLAWVPGELKLVTAAGDQAKFDWKAG---ELIGTC 136
DB 53 SSSDLT-----PVAELGHEGEGVDLSFSPDGRLLASASDDRTVRIWDLAVGGARLVKTL 108
QY 137 KGHQCSLSKVAFSKFEKAVFCTGGRGNIMWDTNCKKDGFRQVNOISGAHNTSDKQT 196
DB 109 TGHNTYAFVCSFSPHGN-VLASGSFDETVRVWEVRSKG-----CLRVLPAHS----- 154

; LOCATION: (1)..(317)
; OTHER INFORMATION: Ceres Seq. ID no. 12724103
US-11-056-355B-94244

Query Match
Best Local Similarity 6.0%; Score 234; DB 7; Length 317;
Matches 75; Conservative 66; Mismatches 126; Indels 64; Gaps 11;

QY 66 VLAVANEGBFVRLY--NTSQSFRKCKFKEMAHNAVFDLAWVPGELKLVTAAGDQAK 123
DB 38 LLASASADKTIIRTYTINTINDPIAEP-VQEFTHENGISDVAFSSDARPIVSADDTLK 96
QY 124 FWDVAGELIGCKGHQCSLSKVAFSKFEKAVFCTGGRGNIMWDTNCKKDGFRQVN 183
DB 97 LMDVETGSLIKTLIGHTNYAFVNFNP-QSNMIVSGSFDETIRIWDVTGK-----CL 148
QY 184 QISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQOSVTVVLFDENTLVSAGAVDGIKVM 243
DB 149 KVLPAHS-----DPTAVDFNRDGLSVSSVSDGLCRIM 182
QY 244 D-----LRKNYATAYROEPIASKSFLYPGSTRKLGYSLLDSTGTLFANCTDDNIYMF 298
DB 183 DSGTGHCVKTLIDDENPPVSFVR-----SPNGKFIILVGTLDNTLRIM 225
QY 299 NMTGLKTSVPAFNCHQNSTFYVKSSLS-PDQFLVSGSSDEAAIYIKVSTPWQPTVLL 357
DB 226 NTSSAKF--LKTYTGHVNAQYCISSAFSVTNGKRIVSGSEDNCVHMWELNSK-KLLQKLE 282
QY 358 GHSQEVTSVCMWPSDFTKIATCDDNTLKIW 388
DB 283 GHTETVMNVACHPTE-NLIASGLDKTVRIW 312

RESULT 14
US-11-056-355B-13096
; Sequence 13096, Application US/11056355B
; Publication No. US20060150283A1
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: 2005-02-14
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 13096
; LENGTH: 346
; TYPE: prt
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(346)
; OTHER INFORMATION: Ceres Seq. ID no. 12344625
US-11-056-355B-13096

Query Match
Best Local Similarity 6.0%; Score 231.5; DB 7; Length 346;
Matches 84; Conservative 63; Mismatches 142; Indels 93; Gaps 14;

QY 20 SSQYPLQSLITGYQCSGNDEHTSYGETGVPPVPPGCTFSSAPNMEHVLAVANEGBFVRLY 79
DB 39 SPGYVLRSTLEGR-----RAVSTVKFSPDG-----RLLASASADKLLRW 79
QY 80 NTESQSRKCKFKEMAHNAVFDLAWVPGELKLVTAAGDQAKFDWKAG---ELIGTC 136
DB 80 SSSDLT-----PVAELGHEGEGVDLSFSPDGRLLASASDDRTVRIWDLAVGGARLVKTL 135
QY 137 KGHQCSLSKVAFSKFEKAVFCTGGRGNIMWDTNCKKDGFRQVNOISGAHNTSDKQT 196
DB 136 TGHNTYAFVCSFSPHGN-VLASGSFDETVRVWEVRSKG-----CLRVLPAHS----- 181
QY 197 PSKPKKQNSKGLAPSVDFQOSVTVVLFDENTLVSAGAVDGIKVMWDLR-----KNYTA 251
DB 182 -----EPVTAVDFDREGDMIVSGYDGLCRVWDSATGCHCVKTLID 221
QY 252 YRQEPFASKSFLYPGSTRKLGYSLLDSTGTLFANCTDDNIYMFNMTG---LKTSPV 308
DB 222 DESPPVSPAKF-----SPNGKFIILATLDTLRLNFSAGKFLKT--- 261
QY 309 AIFNCHQNSTFYVKSSLS-PDQFLVSGSSDEAAIYIKVSTPWQPTVLLGHQSQEVTSV- 366
DB 262 --YTHVNTKVCIPAAFSITNSKYIVSGSEDKCVYLDLQSR-RIVQKLEGHTDITVIAVS 318
QY 367 CWCPSDFTKIATCDDNTLKIW 388
DB 292 CHPKENMIASGALDNDKTVKM 313
```

Db 319 CPKENMIASGALDNDRKTKVW 340

RESULT 15

US-10-449-902-42332

; Sequence 42332, Application US/10449902

; Publication No. US20060123505A1

; GENERAL INFORMATION:

; APPLICANT: National Institute of Agrobiological Sciences.

; APPLICANT: Bio-oriented Technology Research Advancement Institution.

; APPLICANT: The Institute of Physical and Chemical Research.

; APPLICANT: Foundation for Advancement of International Science.

; TITLE OF INVENTION: FULL-LENGTH PLANT CDNA AND USES THEREOF

; FILE REFERENCE: MOA-A0205Y1-US

; CURRENT APPLICATION NUMBER: US/10/449,902

; CURRENT FILING DATE: 2003-05-29

; PRIOR APPLICATION NUMBER: JP 2002-203269

; PRIOR FILING DATE: 2002-05-30

; PRIOR APPLICATION NUMBER: JP 2002-383870

; PRIOR FILING DATE: 2002-12-11

; NUMBER OF SEQ ID NOS: 56791

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 42332

; LENGTH: 1281

; TYPE: PRT

; ORGANISM: Oryza sativa

US-10-449-902-42332

Query Match 6.0%; Score 231.5; DB 6; Length 1281;

Best Local Similarity 23.8%; Pred. No. 5.2e-05;

Matches 98; Conservative 60; Mismatches 143; Indels 111; Gaps 20;

QY 72 EEGFVRLYNTESQSRKCKFKEMAHNAVFDLAWVPGELKLVTAAGDQTAKEWVKAGR 131

Db 625 QEGRVWVENRVEYLSPTQTLRGHTGPFVGTYSADGQRLVSGSDLSRVVWDASTGE 684

QY 132 LIGTCKGHCQSLKSAFSEKAVFTCTGGRDGNIMWMDT---RCNKK---DGFYRQV-- 182

Db 685 SLRELKGTADIRAVALSADGQRI-ASSGDDQTVRVWDASTGECLELKGHTGWRVAI 743

QY 183 -----NOISGAHNTS-----DKQTPSKPKKQNSKGL-----APSVDFQQSVT----- 220

Db 744 SADGQRVVSGSYDQTLRVWMDATGECVRELQGHSTSLVFAVALSADGQRIVSGSSDLTARV 803

QY 221 -----VVLFDENTLVSAVADGIIKWDLRKNYTAHQEPIASK 260

Db 804 WDTATGETLRELKGTGWRVSAFSTDGQRIVTGGDDQSVRVWD-----AS- 849

QY 261 SFLYPGSSSTRKL-GYS-----SLILDSTGTLFANCTDDNIYMFN-MTG-----LKTSPVA 309

Db 850 -----TGECVRELKGYTAALISVAFSPDQRIVSGGGDQTVRVVWNAATGECQCELK----- 900

QY 310 IFNGHONSTFYVKS-SLSPDDQFLVSGSSDEAAYIWKVSTPWQPPT-----VLLG 358

Db 901 ---GH---TEQVDSIAFSPDGHVSGSIDQTLRVWDVSSSLSPSSSSSGGAGLRERQ 954

QY 359 HSGEIVTVCWPCSPDFTKIATCSDNDNLKIWRNRG--LEEKPGCDKLSTVGW 408

Db 955 HTKDVSNAF-PPDGKRLASGDDQSVRVWDVAVSGELLHELQHS-----GW 1000

Search completed: November 8, 2006, 18:19:38

Job time : 45 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2006, 18:14:54 ; Search time 183 Seconds
(without alignments)
1847.797 Million cell updates/sec

Title: US-10-726-160-2
Perfect score: 3877
Sequence: 1 MLFNSVLRQPQLGVLRNGWS.....TYFHRKSQDFCGPEHSTEL 730

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA Main:*
- 1: /EMC_Celestra_SIDS3/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 2: /EMC_Celestra_SIDS3/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 3: /EMC_Celestra_SIDS3/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
 - 4: /EMC_Celestra_SIDS3/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
 - 5: /EMC_Celestra_SIDS3/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
 - 6: /EMC_Celestra_SIDS3/ptodata/2/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3877	100.0	730	4	US-10-145-396-12
2	3877	100.0	730	4	US-10-409-511-2
3	3877	100.0	730	5	US-10-726-160-2
4	3873	99.9	730	3	US-09-780-053-2
5	3503.5	90.4	673	5	US-10-330-773-674
6	3479.5	89.7	729	4	US-10-145-396-11
7	2914.5	75.2	784	5	US-10-330-773-671
8	1864	48.1	347	3	US-09-780-053-4
9	1165	30.0	522	4	US-10-403-571-74
10	1084	28.0	206	4	US-10-221-625-24
11	793.5	20.5	769	6	US-11-097-143-23652
12	759	19.6	146	5	US-10-330-773-676
13	753	19.4	351	3	US-09-780-053-5
14	587	15.1	525	4	US-10-739-930-8452
15	575.5	14.8	555	4	US-10-437-963-118806
16	460.5	11.9	314	4	US-10-424-599-282884
17	406	10.5	111	4	US-10-106-698-4768
18	312	8.0	610	4	US-10-369-493-20185
19	298	7.7	1005	4	US-10-369-493-20242
20	295.5	7.6	579	4	US-10-369-493-18893
21	281	7.2	1595	4	US-10-425-115-205747
22	279.5	7.2	1118	4	US-10-369-493-18845
23	269	6.9	598	4	US-10-425-115-221471
24	269	6.9	608	4	US-10-369-493-20224
25	267	6.9	1155	4	US-10-369-493-19869
26	266	6.9	606	4	US-10-369-493-18907
27	261	6.7	478	4	US-10-369-493-18905

28	259.5	6.7	1356	4	US-10-077-111-10	Sequence 10, Appl
29	258.5	6.7	1117	4	US-10-369-493-18965	Sequence 18965, A
30	256.5	6.6	343	4	US-10-119-932-5	Sequence 5, Appli
31	256	6.6	348	6	US-11-024-959-479	Sequence 479, App
32	256	6.6	515	4	US-10-369-493-1531	Sequence 1531, Ap
33	255.5	6.6	655	4	US-10-369-493-20066	Sequence 20066, A
34	254.5	6.6	1136	4	US-10-369-493-19046	Sequence 19046, A
35	254	6.6	379	4	US-10-369-493-8153	Sequence 8153, Ap
36	254	6.6	742	4	US-10-077-111-11	Sequence 11, Appl
37	253	6.5	375	4	US-10-119-932-1	Sequence 1, Appli
38	253	6.4	415	4	US-10-108-260A-4324	Sequence 4324, Ap
39	250	6.4	1140	4	US-10-369-493-19053	Sequence 19053, A
40	249.5	6.4	347	6	US-11-097-143-6498	Sequence 6498, Ap
41	249	6.4	329	4	US-10-424-599-272779	Sequence 272779, A
42	247.5	6.4	429	4	US-10-369-493-11348	Sequence 11348, A
43	247	6.4	559	4	US-10-369-493-20028	Sequence 20028, A
44	245	6.3	347	4	US-10-425-114-62553	Sequence 62553, A
45	244.5	6.3	626	4	US-10-369-493-20247	Sequence 20247, A

ALIGNMENTS

RESULT 1
US-10-145-396-12
; Sequence 12, Application US/10145396
; Publication No. US20030003485A1
; GENERAL INFORMATION:
; APPLICANT: Uenaka, Akiko
; TITLE OF INVENTION: METHODS FOR IDENTIFYING ANTIGENS
; FILE REFERENCE: L00461/70129
; CURRENT APPLICATION NUMBER: US/10/145,396
; CURRENT FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 60/291,125
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-145-396-12

Query Match	100.0%;	Score 3877;	DB 4;	Length 730;
Best Local Similarity	100.0%;	Pred. No. 1.1e-259;	Mismatches 0;	Indels 0; Gaps 0;
Matches	730;	Conservative 0;		
Qy	1	MLFNSVLRQPQLGVLRNGWSQYPLQSLITGYQCSGNDEHTSYGETGVPPVPGCTFSSA	60	
Db	1	MLFNSVLRQPQLGVLRNGWSQYPLQSLITGYQCSGNDEHTSYGETGVPPVPGCTFSSA	60	
Qy	61	PMMEHLAVANBEGFVRLYNTESQSFRRKKCFKEMAHNAVFDLAWVPGELKLVTAAGDQ	120	
Db	61	PMMEHLAVANBEGFVRLYNTESQSFRRKKCFKEMAHNAVFDLAWVPGELKLVTAAGDQ	120	
Qy	121	TAKFVDVXAGELIGTCKGHQCSLKSAVSKFAVFCGGRDGNIMWDTTRCNKXKDGFR	180	
Db	121	TAKFVDVXAGELIGTCKGHQCSLKSAVSKFAVFCGGRDGNIMWDTTRCNKXKDGFR	180	
Qy	181	QVNQISGAHNTSDKOTPSKPKKKQNSKGLAPSVDPQQSVTVVLFDENTLVISAGAVDGI	240	
Db	181	QVNQISGAHNTSDKOTPSKPKKKQNSKGLAPSVDPQQSVTVVLFDENTLVISAGAVDGI	240	
Qy	241	KVWDLRKNYAYRQBPFIASKSLYPGSSSTRKLGYSLLDSTGFSTLFDNCTDNDIYMFNN	300	
Db	241	KVWDLRKNYAYRQBPFIASKSLYPGSSSTRKLGYSLLDSTGFSTLFDNCTDNDIYMFNN	300	
Qy	301	TGLKTSVPAIFNGHONSTFYVKSSLPDDQFLVSGSSDEAAIYKIVSTPWQPTVLLGHS	360	
Db	301	TGLKTSVPAIFNGHONSTFYVKSSLPDDQFLVSGSSDEAAIYKIVSTPWQPTVLLGHS	360	
Qy	361	QEVTSVCWCPDFTKIATCDDNTLKIWLNRGLSEKPGDDKLSTVGWASQKKSRPGL	420	

```
Db 361 QEVTSVWCWCPDFTKIATCSDNTLKIWLNRGLBEKPGGDKLSTVGWASQKKESRRGL 420
Qy 421 VVTSSQSTPAKAPRVKCNPNSSPSSAACAPSCAGDLPLPNTPTFTSIKTSPAKARSP 480
Db 421 VVTSSQSTPAKAPRVKCNPNSSPSSAACAPSCAGDLPLPNTPTFTSIKTSPAKARSP 480
Qy 481 NRRGSVSSVPKPPSPFKMSIRNWVTRTPSSPPITPPASETKIMSPRKALIPVQKSSQ 540
Db 481 NRRGSVSSVPKPPSPFKMSIRNWVTRTPSSPPITPPASETKIMSPRKALIPVQKSSQ 540
Qy 541 AEACSESRNRVRRRLDSSCLESVKQKVCVSCNCVTELDQVENLHLDLCLLAGNQEDLSK 600
Db 541 AEACSESRNRVRRRLDSSCLESVKQKVCVSCNCVTELDQVENLHLDLCLLAGNQEDLSK 600
Qy 601 DSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLPLRPCGEGSEMVGKENSSENK 660
Db 601 DSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLPLRPCGEGSEMVGKENSSENK 660
Qy 661 NWLLAMAAKRAENPSPSPQTPNSRRQSGKTLPSVPTITPSSMRKICTYFHRKSQED 720
Db 661 NWLLAMAAKRAENPSPSPQTPNSRRQSGKTLPSVPTITPSSMRKICTYFHRKSQED 720
Qy 721 FCGPEHSTEL 730
Db 721 FCGPEHSTEL 730
```

RESULT 2

```
US-10-409-511-2
; Sequence 2, Application US/10409511
; Publication No. US20030162268A1
; GENERAL INFORMATION:
; APPLICANT: IP, Nancy Y
; TITLE OF INVENTION: No. US20030162268A1e1 Gene
; FILE REFERENCE: M98/0553/US
; CURRENT APPLICATION NUMBER: US/10/409,511
; CURRENT FILING DATE: 2003-04-08
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-409-511-2
```

```
Query Match 100.0%; Score 3877; DB 4; Length 730;
Best Local Similarity 100.0%; Pred. No. 1.1e-259;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 MLFNSVLRQPQLGVLRNGWSSQYPLQSLLTGYQCSGNDEHTSYGETGVPPVPPFGCTFSSA 60
Db 1 MLFNSVLRQPQLGVLRNGWSSQYPLQSLLTGYQCSGNDEHTSYGETGVPPVPPFGCTFSSA 60
Qy 61 PMEHVLA VANE BEGFVRLYNTESQSFRRKCKFEKMAHNAVFDLAWVPGLKLVTAAGDQ 120
Db 61 PMEHVLA VANE BEGFVRLYNTESQSFRRKCKFEKMAHNAVFDLAWVPGLKLVTAAGDQ 120
Qy 121 TAKFWDVKAGELIGTCKGHQCSLSVAFSKFEKAVFCTGGRDGNIMWDTNCKNKGDFYR 180
Db 121 TAKFWDVKAGELIGTCKGHQCSLSVAFSKFEKAVFCTGGRDGNIMWDTNCKNKGDFYR 180
Qy 181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQSVTVVLFDENTLVLSAGAVDGI 240
Db 181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQSVTVVLFDENTLVLSAGAVDGI 240
Qy 241 KWDLRKNTAYRQEP IASKSFLYPGSSTRKLGYSLLDSTGSLFANCTDDNIYMFNM 300
Db 241 KWDLRKNTAYRQEP IASKSFLYPGSSTRKLGYSLLDSTGSLFANCTDDNIYMFNM 300
Qy 301 TGLKTS PVAIFNGHQNSTFYVKSSI.PDDQFLVSGSSDEAAAYIWKVSTPWQPPTVLLGHS 360
```

```
Db 301 TGLKTS PVAIFNGHQNSTFYVKSSI.PDDQFLVSGSSDEAAAYIWKVSTPWQPPTVLLGHS 360
Qy 361 QEVTSVWCWCPDFTKIATCSDNTLKIWLNRGLBEKPGGDKLSTVGWASQKKESRRGL 420
Db 361 QEVTSVWCWCPDFTKIATCSDNTLKIWLNRGLBEKPGGDKLSTVGWASQKKESRRGL 420
Qy 421 VVTSSQSTPAKAPRVKCNPNSSPSSAACAPSCAGDLPLPNTPTFTSIKTSPAKARSP 480
Db 421 VVTSSQSTPAKAPRVKCNPNSSPSSAACAPSCAGDLPLPNTPTFTSIKTSPAKARSP 480
Qy 481 NRRGSVSSVPKPPSPFKMSIRNWVTRTPSSPPITPPASETKIMSPRKALIPVQKSSQ 540
Db 481 NRRGSVSSVPKPPSPFKMSIRNWVTRTPSSPPITPPASETKIMSPRKALIPVQKSSQ 540
Qy 541 AEACSESRNRVRRRLDSSCLESVKQKVCVSCNCVTELDQVENLHLDLCLLAGNQEDLSK 600
Db 541 AEACSESRNRVRRRLDSSCLESVKQKVCVSCNCVTELDQVENLHLDLCLLAGNQEDLSK 600
Qy 601 DSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLPLRPCGEGSEMVGKENSSENK 660
Db 601 DSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLPLRPCGEGSEMVGKENSSENK 660
Qy 661 NWLLAMAAKRAENPSPSPQTPNSRRQSGKTLPSVPTITPSSMRKICTYFHRKSQED 720
Db 661 NWLLAMAAKRAENPSPSPQTPNSRRQSGKTLPSVPTITPSSMRKICTYFHRKSQED 720
Qy 721 FCGPEHSTEL 730
Db 721 FCGPEHSTEL 730
```

RESULT 3

```
US-10-726-160-2
; Sequence 2, Application US/10726160
; Publication No. US20050037372A1
; GENERAL INFORMATION:
; APPLICANT: IP, Nancy Y
; APPLICANT: Cheung, William M W
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A RETINOIC ACID REGULATED PROTEIN
; FILE REFERENCE: FP4220B
; CURRENT APPLICATION NUMBER: US/10/726,160
; CURRENT FILING DATE: 2003-12-02
; PRIOR APPLICATION NUMBER: US 10/409,511
; PRIOR FILING DATE: 2003-04-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-726-160-2
```

```
Query Match 100.0%; Score 3877; DB 5; Length 730;
Best Local Similarity 100.0%; Pred. No. 1.1e-259;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 MLFNSVLRQPQLGVLRNGWSSQYPLQSLLTGYQCSGNDEHTSYGETGVPPVPPFGCTFSSA 60
Db 1 MLFNSVLRQPQLGVLRNGWSSQYPLQSLLTGYQCSGNDEHTSYGETGVPPVPPFGCTFSSA 60
Qy 61 PMEHVLA VANE BEGFVRLYNTESQSFRRKCKFEKMAHNAVFDLAWVPGLKLVTAAGDQ 120
Db 61 PMEHVLA VANE BEGFVRLYNTESQSFRRKCKFEKMAHNAVFDLAWVPGLKLVTAAGDQ 120
Qy 121 TAKFWDVKAGELIGTCKGHQCSLSVAFSKFEKAVFCTGGRDGNIMWDTNCKNKGDFYR 180
Db 121 TAKFWDVKAGELIGTCKGHQCSLSVAFSKFEKAVFCTGGRDGNIMWDTNCKNKGDFYR 180
Qy 181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQSVTVVLFDENTLVLSAGAVDGI 240
Db 181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQSVTVVLFDENTLVLSAGAVDGI 240
Qy 241 KWDLRKNTAYRQEP IASKSFLYPGSSTRKLGYSLLDSTGSLFANCTDDNIYMFNM 300
```



```

Db      241  KWDLRKNYTAHQEPIASKSPFLPGSTRKLGYSLLDSTGSTLFANCTDDNIYFMNM 300
Qy      301  TGLKTSFVAIFNGHONSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
Db      301  TGLKTSFVAIFNGHONSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
Qy      361  QEVTSVCWCPSTFKIATCSDNTLKIWRLNRGLEEKPGDKLSTVGMASQKKESRPG 420
Db      361  QEVTSVCWCPSTFKIATCSDNTLKIWRLNRGLEEKPGDKLSTVGMASQKKESRPG 420
Qy      421  VVTSSQSTPAKAPRVKCNPSNPSAAACAPSCAGDLPLPSNTPFTSIKTSAPAKARSP 480
Db      421  VVTSSQSTPAKAPRVKCNPSNPSAAACAPSCAGDLPLPSNTPFTSIKTSAPAKARSP 480
Qy      481  NRGVSVSVPKPPSSFKMSIRNWTRTPSSSPPIPPASETKIMSPRKLIPVSKSSQ 540
Db      481  NRGVSVSVPKPPSSFKMSIRNWTRTPSSSPPIPPASETKIMSPRKLIPVSKSSQ 540
Qy      541  AEACSESRNRVKRRLDSSCLSVKQKVCVSCNCTVTELDGOVENLHLDLCLAGNEDLSK 600
Db      541  AEACSESRNRVKRRLDSSCLSVKQKVCVSCNCTVTELDGOVENLHLDLCLAGNEDLSK 600
Qy      601  DSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTPLPLRPGCGEGSEMVGVKENSSENK 660
Db      601  DSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTPLPLRPGCGEGSEMVGVKENSSENK 660
Qy      661  NWLLAAAKRAENPRSPSSQTPNSRQSGKTLPSVTTITPSSMRKICTYFHRKSQED 720
Db      661  NWLLAAAKRAENPRSPSSQTPNSRQSGKTLPSVTTITPSSMRKICTYFHRKSQED 720
Qy      721  FCGPEHSTEL 730
Db      721  FCGPEHSTEL 730

RESULT 4
US-09-780-053-2
; Sequence 2, Application US/09780053
; Patent No. US20020102640A1
; GENERAL INFORMATION:
; APPLICANT: Rene S. Hubert
; APPLICANT: Daniel E. H. Afar
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Mary Paris
; APPLICANT: Eliana Levin
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 83PSG4: A TISSUE SPECIFIC PROTEIN
; TITLE OF INVENTION: HIGHLY EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: 129.5USU1
; CURRENT APPLICATION NUMBER: US/09/780, 053
; CURRENT FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2001-02-09
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 716
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-780-053-2

```

```

Query Match      99.9%; Score 3873; DB 3; Length 730;
Best Local Similarity 99.9%; Pred. No. 2.1e-259;
Matches 729; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1  MLFNSVLRQPOLGVLNRNGSSQYPLQSLITGYQCSGNDEHTSYGETGVVPPPGCTFSSA 60
Db      1  MLFNSALRQPOLGVLNRNGSSQYPLQSLITGYQCSGNDEHTSYGETGVVPPPGCTFSSA 60
Qy      61  PNMEHLAVANBEGFVRLYNTESQSRKCKFKEMAHNAVFDLAWVPGLKLVTAAGDQ 120

```

```

Db      61  PNMEHLAVANBEGFVRLYNTESQSRKCKFKEMAHNAVFDLAWVPGLKLVTAAGDQ 120
Qy      121  TAKFWDVWVAGELIGTCGHQCSLSKSAFSPKFAVCTGGRDGNIMVWDTRCNKKDGFYR 180
Db      121  TAKFWDVWVAGELIGTCGHQCSLSKSAFSPKFAVCTGGRDGNIMVWDTRCNKKDGFYR 180
Qy      181  QVNIQISGAHNTSDKOTPSKPKKKQNSKGLAPSVDFQQSVTVVLFQDENTLVLSAGAVDGI 240
Db      181  QVNIQISGAHNTSDKOTPSKPKKKQNSKGLAPSVDFQQSVTVVLFQDENTLVLSAGAVDGI 240
Qy      241  KWDLRKNYTAHQEPIASKSPFLPGSTRKLGYSLLDSTGSTLFANCTDDNIYFMNM 300
Db      241  KWDLRKNYTAHQEPIASKSPFLPGSTRKLGYSLLDSTGSTLFANCTDDNIYFMNM 300
Qy      301  TGLKTSFVAIFNGHONSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
Db      301  TGLKTSFVAIFNGHONSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
Qy      361  QEVTSVCWCPSTFKIATCSDNTLKIWRLNRGLEEKPGDKLSTVGMASQKKESRPG 420
Db      361  QEVTSVCWCPSTFKIATCSDNTLKIWRLNRGLEEKPGDKLSTVGMASQKKESRPG 420
Qy      421  VVTSSQSTPAKAPRVKCNPSNPSAAACAPSCAGDLPLPSNTPFTSIKTSAPAKARSP 480
Db      421  VVTSSQSTPAKAPRVKCNPSNPSAAACAPSCAGDLPLPSNTPFTSIKTSAPAKARSP 480
Qy      481  NRGVSVSVPKPPSSFKMSIRNWTRTPSSSPPIPPASETKIMSPRKLIPVSKSSQ 540
Db      481  NRGVSVSVPKPPSSFKMSIRNWTRTPSSSPPIPPASETKIMSPRKLIPVSKSSQ 540
Qy      541  AEACSESRNRVKRRLDSSCLSVKQKVCVSCNCTVTELDGOVENLHLDLCLAGNEDLSK 600
Db      541  AEACSESRNRVKRRLDSSCLSVKQKVCVSCNCTVTELDGOVENLHLDLCLAGNEDLSK 600
Qy      601  DSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTPLPLRPGCGEGSEMVGVKENSSENK 660
Db      601  DSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTPLPLRPGCGEGSEMVGVKENSSENK 660
Qy      661  NWLLAAAKRAENPRSPSSQTPNSRQSGKTLPSVTTITPSSMRKICTYFHRKSQED 720
Db      661  NWLLAAAKRAENPRSPSSQTPNSRQSGKTLPSVTTITPSSMRKICTYFHRKSQED 720
Qy      721  FCGPEHSTEL 730
Db      721  FCGPEHSTEL 730

RESULT 5
US-10-330-773-674
; Sequence 674, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 674
; LENGTH: 673
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-330-773-674

```

```

Query Match      90.4%; Score 3503.5; DB 5; Length 673;
Best Local Similarity 91.8%; Pred. No. 7.4e-234;
Matches 670; Conservative 1; Mismatches 2; Indels 57; Gaps 2;

Qy      1  MLFNSVLRQPOLGVLNRNGSSQYPLQSLITGYQCSGNDEHTSYGETGVVPPPGCTFSSA 60
Db      1  MLFNSVLRQPOLGVLNRNGSSQYPLQSLITGYQCSGNDEHTSYGETGVVPPPGCTFSSA 60

```

```
QY 61 PNMEHLAVANBEGFVRLYNTESQSRKCKCFKEMAHNAVFDLAWVPGELKLVTAAGDQ 120
DB 61 PNMEHLAVANBEGFVRLYNTESQSRKCKCFKEMAHNAVFDLAWVPGELKLVTAAGDQ 120
QY 121 TAKFWDVKAAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIMVWDTNCKKDGFR 180
DB 121 TAKFWDVKAAGELIGTCKGHQCSLKSVAFSKFEK 156
QY 181 QVNIQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQOSVTVVLFDQENTLVSAGAVDGI 240
DB 157 -----DFQOSVTVVLFDQENTLVSAGAVDGI 183
QY 241 KWDLRKNYTAIROEPIASKFELYPGSTRKLGYSLLILDSTGTLFANCTDDNIYMFNM 300
DB 184 KWDLRKNYTAIROEPIASKFELYPGSTRKLGYSLLILDSTGTLFANCTDDNIYMFNM 243
QY 301 TGLKTSFPAIENGHQNSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
DB 244 TGLKTSFPAIENGHQNSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 303
QY 361 QEVTSVCMCPDFTKIATCSDDNITLKIWRLNRGLEEKPGDKLSTVGWASQKKESRPG 420
DB 304 QEVTSVCMCPDFTKIATCSDDNITLKIWRLNRGLEEKPGDKLSTVGWASQKKESRPG 363
QY 421 VVTSSQSTPAKAPRVKCNPSNSPSSAAACAPSCAGDLPLPNTPTFSIKTSPAKARSP 480
DB 364 VVTSSQSTPAKAPRVKCNPSNSPSSAAACAPSCAGDLPLPNTPTFSIKTSPAKARSP 423
QY 481 NRRGVSVPKPPSPSKFMSIRNWTPTSSPPITPPASETKIMSPKALIPVSQKSSQ 540
DB 424 NRRGVSVPKPPSPSKFMSIRNWTPTSSPPITPPASETKIMSPKALIPVSQKSSQ 483
QY 541 AEACSESRNRVRRRLDSSCLSVKQKVCSCNCTELDQVENLHLDLCLLAGNQEDLSK 600
DB 484 AEACSESRNRVRRRLDSSCLSVKQKVCSCNCTELDQVENLHLDLCLLAGNQEDLSK 543
QY 601 DSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLRPGEGSEMVGVKENSSENK 660
DB 544 DSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLRPGEGSEMVGVKENSSENK 603
QY 661 NWLLAMAAKRAENPSPRSPSSQTPNSRRSQSGKTLPSVTTIPSSMRKICTYFHRKSQED 720
DB 604 NWLLAMAAKRAENPSPRSPSSQTPNSRRSQSGKTLPSVTTIPSSMRKICTYFHRKSQED 663
QY 721 FCGPEHSTEL 730
DB 664 FCGPEHSTEL 673
```

RESULT 6

```
US-10-145-396-11
; Sequence 11, Application US/10145396
; Publication No. US20030003485A1
; GENERAL INFORMATION:
; APPLICANT: Uenaka, Akiko
; APPLICANT: Nakayama, Eiichi
; TITLE OF INVENTION: METHODS FOR IDENTIFYING ANTIGENS
; FILE REFERENCE: L00461/70129
; CURRENT APPLICATION NUMBER: US/10/145,396
; CURRENT FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 60/291,125
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 729
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-145-396-11
```

Query Match 89.7%; Score 3479.5; DB 4; Length 729;
Best Local Similarity 89.2%; Pred. No. 3.8e-232;

```
Matches 651; Conservative 36; Mismatches 42; Indels 1; Gaps 1;
QY 1 MLFNSVLKPOLGVLKRWGSSQYFQSLLTQYQCSGNDHSTSYGETGVVPVPGCTFSSA 60
DB 1 MLFNSVLKPOLGVLKRWGSSSHYPQSLLSYQCNCHDHTSYGETGVVPVPGCTFCTA 60
QY 61 PNMEHLAVANBEGFVRLYNTESQSRKCKCFKEMAHNAVFDLAWVPGELKLVTAAGDQ 120
DB 61 PSMEHLAVANBEGFVRLYNTESQTSKKTCTCFKEMAHNAVFDLAWVPGELKLVTAAGDQ 120
QY 121 TAKFWDVKAAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIMVWDTNCKKDGFR 180
DB 121 TAKFWDVKAAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIMVWDTNCKKDGFR 180
QY 181 QVNIQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQOSVTVVLFDQENTLVSAGAVDGI 240
DB 181 QVNIQISGAHNTADKQTPSKPKKQNSKGLAPSVDFQOSVTVVLFDQENTLVSAGAVDGI 240
QY 241 KWDLRKNYTAIROEPIASKFELYPGSTRKLGYSLLILDSTGTLFANCTDDNIYMFNM 300
DB 241 KWDLRKNYTAIROEPIASKFELYPGSTRKLGYSLLILDSTGTLFANCTDDNIYMFNM 300
QY 301 TGLKTSFPAIENGHQNSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
DB 301 TGLKTSFPAVFNHGHQNSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSMHPHPPTVLLGHS 360
QY 361 QEVTSVCMCPDFTKIATCSDDNITLKIWRLNRGLEEKPGDKLSTVGWASQKKESRPG 420
DB 361 QEVTSVCMCPDFTKIATCSDDNITLKIWRLNRGLEEKPGDKLSTVGWASQKKESRPG 420
QY 421 VVTSSQSTPAKAPRVKCNPSNSPSSAAACAPSCAGDLPLPNTPTFSIKTSPAKARSP 480
DB 420 VVTSSQSTPAKAPRAKSSPISPPSACTPSCAGDLPLPNTPTFSIKTSPATRSV 479
QY 481 NRRGVSVPKPPSPSKFMSIRNWTPTSSPPITPPASETKIMSPKALIPVSQKSSQ 540
DB 480 SRRGSISSVPKPLSSFKMSLRNWTPTSSPPITPPASETKISSPKALIPVSQKSSQ 539
QY 541 AEACSESRNRVRRRLDSSCLSVKQKVCSCNCTELDQVENLHLDLCLLAGNQEDLSK 600
DB 540 ADACSESRNRVRRRLDSSCLSVKQKVCSCNCTELDQVAESLRLDCLLSGTQEVLSQ 599
QY 601 DSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLRPGEGSEMVGVKENSSENK 660
DB 600 DSEGTKSSKIEGAGTSISEPPSPVPSYASGCGPLPLRPGEGSEMVGVKENSSENK 659
QY 661 NWLLAMAAKRAENPSPRSPSSQTPNSRRSQSGKTLPSVTTIPSSMRKICTYFHRKSQED 720
DB 660 NWLLAMAAKRAENPSPRSPSSQTPNSRRSQSGKTLPSVTTIPSSMRKICTYFHRKTQDD 719
QY 721 FCGPEHSTEL 730
DB 720 FCSPEHSTEL 729

RESULT 7
US-10-330-773-671
; Sequence 671, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 671
; LENGTH: 784
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-330-773-671
```

```

Query Match      75.2%; Score 2914.5; DB 5; Length 784;
Best Local Similarity 70.6%; Pred. No. 5.9e-193;
Matches 570; Conservative 34; Mismatches 46; Indels 157; Gaps 5;

QY 1 MLFNSVLRLPOLGVLNR-----17
DB 39 MLFNSVLRLPOLGVLNRGAPPWIAADCGPGLPRPERIRASDRRTTSCLGSRFTLFTVI 98
QY 18 -----GSSQYPLQSLLTGYCSGNDEH 40
DB 99 ELFSRRPKLGRSVASRAPGLRQYRALTRCTTLVRVGGSSHYPLQSLLTGYCSGNDEH 158
QY 41 TSGETGVPPVPPGCTFSSAPNMEHVLAVANEGFVRLYNTESQSPKCKFKEMAHWNA 100
DB 159 TSGETGVPPVPPGCTFCTAPSMEHILAVANEGFVRLYNTESQTSKTKCFK-----210
QY 101 VFDLAWVPGELKLVTAAGDQTAKFMDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFTCG 160
DB 211 -----GG 212
QY 161 RDGNIMWDTRCNKDGFYQVQNOISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQGSVT 220
DB 213 RDGNIMWDTRCNKDGFYQVQNOISGAHNTADKQTPSKPKKQNSKGLAPVDSQGSVT 272
QY 221 VVLFQDENTLVSAVADGIIKWDLRKNYTAQROEPIASKSFYLPGSSSTRKGLGYSSILLD 280
DB 273 VVLFQDENTLVSAVADGIIKWDLRKNYTAQROEPIASKSFYLPGSTSTRKGLGYSSILVD 332
QY 281 STGSTLFPANCTDDNIYMFNMTGLKTSVPAIFNGHQNSTFYVKSLSLSPDDQFLVSGSDEA 340
DB 333 STGSTLFPANCTDDNIYMFNMTGLKTSVPAIFNGHQNSTFYVKSLSLSPDDQFLVSGSDEA 392
QY 341 AYIWKVSTPQPPPTVLLGHGSEVTSVCWCPSPDFTKIATCSDNTLKIWLNRGLGKPKGG 400
DB 393 AYIWKVSTPQPPPTVLLGHGSEVTSVCWCPSPDFTKIATCSDNTLKIWLNRGLGKPK-G 451
QY 401 DKLSTVCMASOKKESRPGVLTVTSSQTPAKAPRVKCNPSNPSGSAACAPSCAGDLPL 460
DB 452 DKHSIVGWTQKKEVKACPVTVFSSQSTPAKAPRAKSPSISPSAACPSCAGDLPL 511
QY 461 PSNTPTFSIKTSPAKARSPINRRGSSVSVPKPPSPFMSIRNWTPTPSSPPITPPAS 520
DB 512 PSSTPTFSVKTTPATTRSSVSRGSISSVSPKPLSLSPFMSLRNWTPTPSSPPVTPPAS 571
QY 521 ETKIMSPRKALIPVQKSSQAEACSESRNRVRRRLDSSCLSVKQKCVKSCNCVTELDGQ 580
DB 572 ETKISSPRKALIPVQKSSQADACSESRNRVRRRLDSSCLSVKQKCVKSCNCVTELDGQ 631
QY 581 VENLHLDLCLAGNOEDLSKDSLGPTKSKIEGAGTSISEPPSPISPYASESCGTLPLPL 640
DB 632 AESLRDLCLLSTGQEVLSQDSGPTKSKTEGAGTSISEPPSPVPSYASEGCGPLPL 691
QY 641 RPOGSEGMVKGKNSPENKNWLLAAKKAENSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 700
DB 692 RPOGSEGMVKGKNSPENKNWLLAAKKAENSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 751
QY 701 ITPSSMRKIC-----TYFHRKSQEDFCG 723
DB 752 LA-----VCGNHQVLVDAGIKFCFCG 772

RESULT 8
US-09-780-053-4
; Sequence 4, Application US/09780053
; Patent No. US20020102640A1
; GENERAL INFORMATION:
; APPLICANT: Rene S. Hubert
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Pia M. Challita-Eid
; APPLICANT: Mary Paris
; APPLICANT: Elana Levin
; APPLICANT: Steve Chappell Mitchell

```

```

; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 83PSG4: A TISSUE SPECIFIC PROTEIN
; FILE REFERENCE: 129.5USU1
; CURRENT APPLICATION NUMBER: US/09/780,053
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/181,261
; NUMBER OF SEQ ID NOS: 716
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Homo Sapiens
; ORGANISM: Homo Sapiens
US-09-780-053-4

Query Match      48.1%; Score 1864; DB 3; Length 347;
Best Local Similarity 100.0%; Pred. No. 1.2e-120;
Matches 347; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 PVPFGCTFSSAPNMEHVLAVANEGFVRLYNTESQSPKCKFKEMAHWNAVFDLAWVP 108
DB 1 PVPFGCTFSSAPNMEHVLAVANEGFVRLYNTESQSPKCKFKEMAHWNAVFDLAWVP 60
QY 109 GELKLVTAAGDQTAKFMDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFTCGRDNIMVW 168
DB 61 GELKLVTAAGDQTAKFMDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFTCGRDNIMVW 120
QY 169 DTRCNKKGDFYQVQNOISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQGSVTVVLFDEN 228
DB 121 DTRCNKKGDFYQVQNOISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQGSVTVVLFDEN 180
QY 229 TLVSAGAVDGIKWDLRKNYTAQROEPIASKSFYLPGSSSTRKGLGYSSILLDSTGTLFA 288
DB 181 TLVSAGAVDGIKWDLRKNYTAQROEPIASKSFYLPGSSSTRKGLGYSSILLDSTGTLFA 240
QY 289 NCTDDNIYMFNMTGLKTSVPAIFNGHQNSTFYVKSLSLSPDDQFLVSGSDEAAYIWKVST 348
DB 241 NCTDDNIYMFNMTGLKTSVPAIFNGHQNSTFYVKSLSLSPDDQFLVSGSDEAAYIWKVST 300
QY 349 PQPPTVLLGHGSEVTSVCWCPSPDFTKIATCSDNTLKIWLNRGLG 395
DB 301 PQPPTVLLGHGSEVTSVCWCPSPDFTKIATCSDNTLKIWLNRGLG 347

RESULT 9
US-10-403-571-74
; Sequence 74, Application US/10403571
; Publication No. US20040068763A1
; GENERAL INFORMATION:
; APPLICANT: Hopkins, Nancy
; APPLICANT: Golling, Gregory
; APPLICANT: Amsterdam, Adam
; APPLICANT: Sun, Zhaoxia
; TITLE OF INVENTION: Developmental Mutations in Zebrafish
; FILE REFERENCE: 01997/539002
; CURRENT APPLICATION NUMBER: US/10/403,571
; CURRENT FILING DATE: 2003-03-25
; PRIOR FILING DATE: 2002-03-29
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Danio rerio
US-10-403-571-74

Query Match      30.0%; Score 1165; DB 4; Length 522;
Best Local Similarity 38.1%; Pred. No. 5.5e-72;
Matches 275; Conservative 74; Mismatches 139; Indels 234; Gaps 20;

QY 16 RNCWSSQYPLQSLLTGYQCSGNDEHTSYGETGVPPFPFGCTFSSAPNMEHVLAVANEGF 75

```

```
Db      17  RINGEORLFLPSLLDCECARDEHISYGASAAVPPPGCTFSSAHGQONCLAVANEGBF 76
Qy      76  VRLYNT-ESQSPRKKCFKEMAHWNAVDFLAWVPGELKLVTAAAGDQTAKEFMDVKAGELIG 134
Db      77  VTIFNTGKQS---SVLKEMQAHNAVDIAWVPGTNCCLVTASGDTARLWDVITGDLG 133
Qy     135  TKGHQCSLSKVAFSKFAVCTGGRDGNIMVWDTNCKKDGFRQVNOISGAHNTSDK 194
Db     134  TFKGHQCSLSKVAFYKQKAVFSTGGRDGNIMVWDTNCKSKDGFRQVNOISGAHMKPER 193
Qy     195  QTPSPKPKQNSKGLAPSVDFOGVTVLVDFQDENTLIVSAGAVDGIKWDLRKNTAYKRO 254
Db     194  FTPQTKR-----RGMAPPVDSQGVTVLDFDETKLISGAGVDGIKWDLRNTAYHQ 249
Qy     255  EPIASKSFLPGSSTRKGLGYSSILLDSTGTSTLPANCTDDNIYFMNMTGLKTPVAIFNGH 314
Db     250  NPLPQAYPYPSCTKGLYSGLSLDYTGSRLFNSCTDDNIYFNISGLKTTPIAVFSGH 309
Qy     315  QNSTFYVKSLSGDDQDFLVSGSSDEAAYIWKVSTPWPQPTVLGHSGQVTSVCMQSPDPT 374
Db     310  SNSSFYVKSIVSPD-----SYRCRFTPP-----IKH-----LFCTS--- 340
Qy     375  KIATCDDNTLKIWLNRGLBEPKGGDKLSTVGWASQKKKESRPGLVTVTSSQSTPAKAP 434
Db     341  -----SOAHOPQNAIIPSTMDPQQ--- 360
Qy     435  RVKCNPSNPSSSAAACAPSCAGDLPLPSNTPTFTSIKTPAKARSPINRRGVSVSVPKPP 494
Db     361  QISC-----SKPHTCSP-----GS----- 375
Qy     495  S9FKMSIRNWNVTTPSSSPPIPTPASETKIMSPRKALIPVSKQSQAECSSSRNRVKRR 554
Db     376  -----VLRTPCQTP-----TGDRFCQLWFGRGDRR 401
Qy     555  LDSSCLESVKQKVCSCNVTDLGOVENIHLDLCLAGNQEDLSKDSLGPTKSSKIEGA 614
Db     402  -----CVLYPN-VKRSRSSVSTLK-----KEDSFG-LESEKQGS 434
Qy     615  GTSISPPSPISPYASESCGTLPLPLRPGEGSEMVKGKNSPENKNWLLAMAARKAEN 674
Db     435  -----DGAESGKNSPPRTDLSVISQKFKG-S 463
Qy     675  P9PRSPSOTPNRRSGKTLSPVITP-----SSMRKICTYFHRKS 717
Db     464  AQPKSPSS---GSSQODTRLESAAVSPRPMKVFSPTNKKASPSKPMKISSYPMKRT 520
Qy     718  QE 719
Db     521  QD 522
```

RESULT 10

```
US-10-221-625-24
; Sequence 24, Application US/10221625
; Publication No. US20040033942A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: YOE, Henry
; APPLICANT: LAL, Preeti
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BANDMAN, Olga
; APPLICANT: TANG, Y. Tom
; APPLICANT: MATHUR, Preete
; APPLICANT: SHAH, Purvi
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: REDDY, Roopa
; TITLE OF INVENTION: TRANSCRIPTION FACTORS
; FILE REFERENCE: PF-0761 PCT
```

```
; CURRENT APPLICATION NUMBER: US/10/221,625
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PERL Program
; SEQ ID NO 24
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040033942A1 1377380CD1
US-10-221-625-24
```

```
Query Match      28.0%; Score 1084; DB 4; Length 206;
Best Local Similarity 100.0%; Pred. No. 6.4e-67;
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      525  MSPRKALIPVSKQSQAECSESENVRKRRDSSCLESVKQKVCSCNVTDLGOVENL 584
Db      1  MSPRKALIPVSKQSQAECSESENVRKRRDSSCLESVKQKVCSCNVTDLGOVENL 60
Qy     585  HLDLCLAGNQEDLSKDSLGPTKSKIEGAGTTSISEPPSPISPYASESCGTLPLRPGC 644
Db      61  HLDLCLAGNQEDLSKDSLGPTKSKIEGAGTTSISEPPSPISPYASESCGTLPLRPGC 120
Qy     645  EGSEMVKGKNSPENKNWLLAMAARKAENSPRSPSSOTPNRRSGKTLSPVITPS 704
Db     121  EGSEMVKGKNSPENKNWLLAMAARKAENSPRSPSSOTPNRRSGKTLSPVITPS 180
Qy     705  SMRKICTYFHRKSQEDFCGPEHSTEL 730
Db     181  SMRKICTYFHRKSQEDFCGPEHSTEL 206
```

RESULT 11

```
US-11-097-143-23652
; Sequence 23652, Application US/11097143
; Publication No. US2005020858A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CLO00728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23652
; LENGTH: 769
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-23652

Query Match      20.5%; Score 793.5; DB 6; Length 769;
Best Local Similarity 29.2%; Pred. No. 5.2e-46;
```

```

Matches 224; Conservative 112; Mismatches 268; Indels 163; Gaps 25;
QY 49 PVPP-FGCTFSSAPNMEHVLAVANESGFVRLYNT-----ESQSF-RKKCFKEMAHWN 99
Db 49 PEPIPSAKFANCDGVRHILAIANEDGKITLQDTTQRNHQPEEQSLVGPQC-----HFN 102
QY 100 AVFDLAWVPGEKLVTAAGDQTAKEFWDVWVAGELIG--TCKGHQCSLKSVAFSKFEKAVFC 157
Db 103 AVFDLEWAPQGMRFVVSASGDHTARLWEVAGSGIRGLNSYVGHTRSVKSAAFKRTDPAVFA 162
QY 158 TGRDGNIMVWDRCNKKGDFYRVQNOISGAHNTSDKQTPSKPKKONSKGLAPSVDFOQ 217
Db 163 TGRDGAILLIDIRANLNMDLTSRVNDCIYSGHTGGTGFVSQKQRTTRPKMAGTTSS 222
QY 218 SVTVLVFODENTLVSAVADGIIKVDLRKNYTAAYRQEPKIASKFLYPGSSSTRKLGYSYL 277
Db 223 SITGLAFQNDNLISCGAGDGVIKVWDLRRNYTAYKKEPLPRHKLPRYAGSSTFR-GFTNL 281
QY 278 ILDTSTGLTFANCTDDNIYMFNMTGLKTPVAIFNGHONSTFYVKSLSLSPDDQFLVSGSS 337
Db 282 IVDSAGTRLYANCMNDTIYCYNLASYQRPLACYKGLNSTFYIKSLSPDGKYLSSGS 341
QY 338 DEAYITWVSTPMQPTVLGHQSQVTSVCMCPDFTKIATCDDNTLKIWRNL-----391
Db 342 DERAYIWNLDHAEPLVALAGHTVEVTCVAMGSSHDCPIVTCSDDHAKIWRIGPDLGL 401
QY 392 -----RGLBEKPGDK-----402
Db 402 SEABRAEKYRGTYASYREFGKKAFGSPSGNHNKLNRLDLESTPKSLKRLMDQNETPGSV 461
QY 403 -----LSTVGWASQ-----KKESRPLGLVTVTSSQSTPAKAPRVK-CNPS 441
Db 462 KITTKESFLEMLGVAQOETEATEQOKRAKPLESRGRRLFGPSSQTAHCHQLQLOINEE 521
QY 442 NSPSSAACAPSCAGDLPLPSNTPTTSIKTSPAKARSPINRRGSSVSSPKPPSPKMSI 501
Db 522 DASPSKRQENSAEDV-----SPLHKLSTP--SHSPLSE--NVNVTSPPTTS-----568
QY 502 RNWVTTPSSPPIITPASETKIMSPRKALIPVSQKSSQAE--ACSESNNVKRRLDSSC 559
Db 569 -----AAAAVAADALNPPISAAIYSPNSL-PNYVLDGEAPHLGIMSPKRAKEKYD--W 622
QY 560 LESVKOKVKSCNCVTELDGQVENHLDLCLAGNOEDLSKDSLGPTKSSKIEGAGTSI 618
Db 623 LNIHQKLSMGRAHVTLEKI-----SEEQADVLA---SPRIQSLRQSEC 666
QY 620 EPPSPISP-----YASESCGTLPLPLRPGCGSEMVGKENSPPENKNWLLAMAARKAEN 674
Db 667 SPRIHASPRRISHTDGGGT-----PAGSSSH--SHSQSQPKTPT-----SSRNSSET 713
QY 675 P-----SPRSPSSQTPNSRRQSGKTLPSVPTI-----TPSSMR 707
Db 714 TLLRPFISIQRSSSVPAEETTTTNAAPSSSDPHPPAVTAPAAATPLSMR 760

RESULT 12
US-10-330-773-676
; Sequence 676, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 676
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

```

```

; NAME/KEY: VARIANT
; LOCATION: (1)....(146)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-330-773-676

Query Match 19.6%; Score 759; DB 5; Length 146;
Best Local Similarity 100.0%; Pred. No. 1.3e-44;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 499 MSIRWVTRTPSSSPPIITPASETKIMSPRKALIPVSQKSSQAEACSESNNVKRRLDSS 558
Db 1 MSIRWVTRTPSSSPPIITPASETKIMSPRKALIPVSQKSSQAEACSESNNVKRRLDSS 60
QY 559 CUESVKOKVKSCNCVTELDGQVENHLDLCLAGNOEDLSKDSLGPTKSSKIEGAGTSI 618
Db 61 CUESVKOKVKSCNCVTELDGQVENHLDLCLAGNOEDLSKDSLGPTKSSKIEGAGTSI 120
QY 619 SEPPSPISPYASESCGTLPLPLRPC 643
Db 121 SEPPSPISPYASESCGTLPLPLRPC 145

RESULT 13
US-09-780-053-5
; Sequence 5, Application US/09780053
; Patent No. US20020102640A1
; GENERAL INFORMATION:
; APPLICANT: Rene S. Hubert
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Pia M. Challita-Bid
; APPLICANT: Mary Faris
; APPLICANT: Elana Levin
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 83P5G4: A TISSUE SPECIFIC PROTEIN
; TITLE OF INVENTION: HIGHLY EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: 129.5USU1
; CURRENT APPLICATION NUMBER: US/09/780,053
; CURRENT FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/181,261
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 716
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-780-053-5

Query Match 19.4%; Score 753; DB 3; Length 351;
Best Local Similarity 42.7%; Pred. No. 1.1e-43;
Matches 153; Conservative 62; Mismatches 125; Indels 18; Gaps 6;

QY 49 PVPP-FGCTFSSAPNMEHVLAVANESGFVRLYNT-----ESQSF-RKKCFKEMAHWN 99
Db 1 PEPIPSAKFANCDGVRHILAIANEDGKITLQDTTQRNHQPEEQSLVGPQC-----HFN 54
QY 100 AVFDLAWVPGEKLVTAAGDQTAKEFWDVWVAGELIG--TCKGHQCSLKSVAFSKFEKAVFC 157
Db 55 AVFDLEWAPQGMRFVVSASGDHTARLWEVAGSGIRGLNSYVGHTRSVKSAAFKRTDPAVFA 114
QY 158 TGRDGNIMVWDRCNKKGDFYRVQNOISGAHNTSDKQTPSKPKKONSKGLAPSVDFOQ 217
Db 115 TGRDGAILLIDIRANLNMDLTSRVNDCIYSGHTGGTGFVSQKQRTTRPKMAGTTSS 174
QY 218 SVTVLVFODENTLVSAVADGIIKVDLRKNYTAAYRQEPKIASKFLYPGSSSTRKLGYSYL 277
Db 175 SITGLAFQNDNLISCGAGDGVIKVWDLRRNYTAYKKEPLPRHKLPRYAGSSTFR-GFTNL 233
QY 278 ILDTSTGLTFANCTDDNIYMFNMTGLKTPVAIFNGHONSTFYVKSLSLSPDDQFLVSGSS 337
Db 234 IVDSAGTRLYANCMNDTIYCYNLASYQRPLACYKGLNSTFYIKSLSPDGKYLSSGS 293

```

```
QY 338 DEAAIYWKVTPWQPPTVLLGHSEVTSVCMCPSTFTKIATCDDNTLKIWRNLGRLE 395
Db 294 DERAYIWNLDHAEPLVALAGHTVEVTCVAGSSHDPCIVTSCDDARHKIMRIGPDLD 351

RESULT 14
US-10-739-930-8452
; Sequence 8452, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 8452
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: ZEAMA-23APR03-C71826_1.p
US-10-739-930-8452

Query Match 15.1%; Score 587; DB 5; Length 523;
Best Local Similarity 31.2%; Pred. No. 6.2e-32;
Matches 158; Conservative 84; Mismatches 194; Indels 70; Gaps 17;

QY 39 EHTSYGETGVPPVPPGCTFSSAPNMEHVLAVANEEGFVRLYNTE-----SQSPKKK--- 89
Db 54 EH-----SGDPAIPFAISFGKTAQTCNLAAVADEGVGLYDTRRLPSASSLSKSAET 108
QY 90 --CFKEWMAHNAVFDLAWVPGELKLVTAAAGQTAKFNDVKAGELIGTCGHCQSLKVA 147
Db 109 RVC--DWVAHNAVFDVCMVKEGSQLTASGQTVKIWSGVGNKRCIGVLSGHTGVKSL 166
QY 148 FSKFEKAVFCTGRDGNIMWTRCNKDGFRQVQVQISGANTSDKQTPSPKPKKQNSK 207
Db 167 CHSSNPFLVSGSRDGSFALWDLRCDPKSPNKGHGETCLTSSAVVREAHSPIL-RSRTRSR 225
QY 208 GLAPSVDFQGSVTVVLF--ODENTLVSAGAVDGLIKVWDLR-----KNYTAVRQEPTA 258
Db 226 AKAAST-----SITSVLKDGVSATSGAADNVVKIWDTRNLKVPFSNKNQSAQAQ-PLE 280
QY 259 SKSFLYPGSSTRKLGYSSILSDSTGLFANCTDDNIYFNMTGLKTSVPAIFNGHQNST 318
Db 281 GV-----KHGISCLSQDSYGAYIAASCWMDNRIYLYSVLVHVNKGPIKVYTGSKIES 330
QY 319 FVYKSSLPDDQFLVSGSSDEAYIWKVSTPWPQPTVLLGHSEVTSVCMCPSTFTKIAT 378
Db 331 FVYKSAISPDGTHILGGSSDGSVYLWQVDPESAPIVLKGHEGEATSDVMWCASEVGKIAT 390
QY 379 CSDDNTLKIWRNLGRLEKPGDKLSTVGWASQ---KKESRPLVTVTSSQSTPAKAPR 435
Db 391 SSDDSTVVRWSTM-----KMCTNVSPFTVRKKAITAPSTEYRSRITHEPATSR 440
QY 436 --VKCNPSNPSPPSSAACAPSCAGDLPLPNTPTFSIKTSPA-----KASPINRRGS 485
Db 441 DGWVCTSGDGLQSGCHSP--LQPRALHFGTPESSKKRALALFEEBALDRSKSPQAQTS 498
QY 486 VSSVSPKPPSPK-MSIRWVWTRTPS 510
Db 499 PSSVL-SPPSSLKRRTIRDYFAAAS 523
```

```
RESULT 15
US-10-437-963-118806
; Sequence 118806, Application US/10437963
; Publication No. US2004012334A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
```

```
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 118806
; LENGTH: 555
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_22082C.1.pep
US-10-437-963-118806
```

```
Query Match 14.8%; Score 575.5; DB 4; Length 555;
Best Local Similarity 29.8%; Pred. No. 4.2e-31;
Matches 162; Conservative 92; Mismatches 200; Indels 89; Gaps 20;

QY 5 SVLRQPQLGVLRNGWSSQYPLQSLLTGYQCSG--NDEHTSYGETGVPPVPPGCTFSSAPN 62
Db 31 SAARLPYLADI-----SSYP-----GGRGSGVIAVEH-----AGDPAIPFAISFCKAEQ 74
QY 63 MEHVLAVANEESGFVRLYNTE-----ESQSPKKK---FKEWMAHNAVFDLAWVPGELKL 113
Db 75 ISHLLAVALADEGVYGIYDTRRLRPSSSSSLKSAETKMSDWAHNAIFDVCWIKDGSQ 134
QY 114 VTAAGDQTAKFNDVKAGELIGTCGHCQSLKSVAFSEKAVFCTGRDGNIMWTRCN 173
Db 135 LTASGDTQTVKIWSVENKCKGLVLSGHTGVKSLSCHSSNPFLIVTGRDGSFALWDLRID 194
QY 174 -KDGFRQVQVQISG-----AHNTSDKQTPSKPKKQNSKGLAPSVDFQGSVTVVLF-QD 226
Db 195 PKTPNGHREACMLGSLVVKQAHSPQ-----RNRTSRKAASST-----SITSVLKLD 243
QY 227 ENTLSAGAVDGLIKVWDLRKNYTAIROEPIASKSFLYPGSSTRKLGYSSILSDSTGL 286
Db 244 DISIATSGAADNIVKIMWTRNLKLSNR--SSQAAMQPLEGV-KHGISCLSQDSYGAYI 300
QY 287 FANCTDDNIYFNMTGLKTSVPAIFNGHQNSTFVYKSSLPDDQFLVSGSSDEAYIWKV 346
Db 301 AASCMDN-----SALHMDKGPIKAYTGSKIESFFYKSAISPDGTHILGGSSDGNVYLWQV 355
QY 347 STPWQPPTVLLGHSEVTSVCMCPSTFTKIATCDDNTLKIWRNLGRLEKPGDKLSTV 406
Db 356 DQPERGPILLGHEGEATSDVMWCASEVGKIATSSDDSKVRVWNTERRV--FPNTSSPTVI 413
QY 407 GWASQKKEPSRPLVTVTSSQSTPAKPRVKCNPSNPSPPSSAACAPSCAGDLPLPNT- 465
Db 414 -----RKRIAPN-----TGRSASHEL-----ATTSRDVGVAACSTADGELPTGSRSP 458
QY 466 -----TPSIKTSAPK-----ARSPINRGSGVSSVSPKPPSPKMSIRWVWTR 507
Db 459 QPRVLEFGTPESSAKRAFLRFQEDSLDIRKSPQAQMSPPSVLSPHSLKRRTIRDYFAS 518
QY 508 TPS 510
Db 519 SSS 521
```

Search completed: November 8, 2006, 18:18:49
Job time : 186 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2006, 18:14:04 ; Search time 53 Seconds
(without alignments)
1205.611 Million cell updates/sec

Title: US-10-726-160-2

Perfect score: 3877

Sequence: 1 MLFNSVLRPQLGVLNGWS.....TYFHRKSQEDFCGPEHSTEL 730

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_AA:*

- 1: /EMC Celerra_SIDS3/ptodata/2/iaa/5 COMB.pcp:*
- 2: /EMC Celerra_SIDS3/ptodata/2/iaa/6 COMB.pcp:*
- 3: /EMC Celerra_SIDS3/ptodata/2/iaa/7 COMB.pcp:*
- 4: /EMC Celerra_SIDS3/ptodata/2/iaa/H_COMB.pcp:*
- 5: /EMC Celerra_SIDS3/ptodata/2/iaa/PCTUS_COMB.pcp:*
- 6: /EMC Celerra_SIDS3/ptodata/2/iaa/RE_COMB.pcp:*
- 7: /EMC Celerra_SIDS3/ptodata/2/iaa/backfiles1.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	256.5	6.6	343	2	US-09-063-743-5
2	256.5	6.6	343	2	US-09-590-540-5
3	253	6.5	375	2	US-09-063-743-1
4	253	6.5	375	2	US-09-590-540-1
5	245.5	6.3	514	1	US-08-190-802A-66
6	245.5	6.3	514	2	US-08-477-346-66
7	245.5	6.3	514	2	US-08-473-089-66
8	245.5	6.3	514	2	US-08-487-072A-66
9	241	6.2	251	2	US-09-291-170A-13
10	241	6.2	251	2	US-09-724-884-13
11	241	6.2	251	2	US-09-724-592-13
12	241	6.2	251	2	US-09-673-222-13
13	233.5	6.0	606	1	US-08-883-534-3
14	233.5	6.0	606	2	US-09-204-764-3
15	233.5	6.0	627	2	US-09-949-016-7082
16	233.5	6.0	627	2	US-09-949-016-7083
17	230.5	5.9	416	2	US-09-270-767-41646
18	211.5	5.5	587	2	US-08-899-578-2
19	209.5	5.4	395	2	US-09-032-372-1
20	208.5	5.4	409	1	US-08-283-917-3
21	208.5	5.4	409	1	US-08-861-716-3
22	208.5	5.4	409	2	US-09-538-092-1119
23	208.5	5.4	410	1	US-08-283-917-9
24	208.5	5.4	410	1	US-08-961-716-9
25	207.5	5.4	690	2	US-09-724-592-2
26	207.5	5.4	690	2	US-09-673-222-2

27	207	5.3	423	2	US-09-013-118-1	Sequence 1, Appli
28	206.5	5.3	690	2	US-09-291-170A-2	Sequence 2, Appli
29	206.5	5.3	690	2	US-09-724-884-2	Sequence 2, Appli
30	201.5	5.2	540	2	US-09-213-888-7	Sequence 7, Appli
31	201.5	5.2	540	2	US-09-213-888-10	Sequence 10, Appli
32	201.5	5.2	540	2	US-09-328-877D-7	Sequence 7, Appli
33	201.5	5.2	540	2	US-09-328-877D-10	Sequence 10, Appli
34	201.5	5.2	545	2	US-09-213-888-6	Sequence 6, Appli
35	201.5	5.2	545	2	US-09-328-877D-6	Sequence 6, Appli
36	201.5	5.2	553	2	US-09-213-888-5	Sequence 5, Appli
37	201.5	5.2	553	2	US-09-328-877D-5	Sequence 5, Appli
38	201.5	5.2	559	2	US-09-213-888-9	Sequence 9, Appli
39	201.5	5.2	559	2	US-09-328-877D-9	Sequence 9, Appli
40	201.5	5.2	589	2	US-09-213-888-8	Sequence 8, Appli
41	201.5	5.2	589	2	US-09-328-877D-8	Sequence 8, Appli
42	201.5	5.2	592	2	US-09-213-888-4	Sequence 4, Appli
43	201.5	5.2	592	2	US-09-328-877D-4	Sequence 4, Appli
44	201.5	5.2	626	2	US-09-213-888-21	Sequence 21, Appli
45	201.5	5.2	626	2	US-09-328-877D-21	Sequence 21, Appli

ALIGNMENTS

RESULT 1

US-09-063-743-5
; Sequence 5, Application US/09063743
; Patent No. 6242214
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Patterson, Chandra
; TITLE OF INVENTION: HUMAN GTPASE-ASSOCIATED PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/063,743
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Carrone, Michael C
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0508 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 343 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 2281093
US-09-063-743-5


```
Query Match 6.6%; Score 256.5; DB 2; Length 343;
Best Local Similarity 25.6%; Pred. No. 1.9e-12;
Matches 79; Conservative 44; Mismatches 117; Indels 69; Gaps 9;

QY 89 KCFKEMAHNAVFDLAWVPGELKLVTAAGDQTAKEFMDVKAGELIGTCKGHQCSLKSVAF 148
DB 87 KFMVLKGHNAILDLHWTSDGSQIVSASPKTKVRAWDVETGKQIKMAEHSFVNSCCP 146
QY 149 SKFEKAVFCTGGRGNIMVWDTRCNKKDGFYRQVNIQISGAHNTSDKQTPSKPKKQNSKG 208
DB 147 TRRGPPLIISGDDGTAKLWDMR-----QRG 172
QY 209 LAPSVDFQSVTVVLVFDQENTLVLSAGAVDGIKWDLRKNYTAHQEPIASKSFYLPSS 268
DB 173 AIQTPDKYQITAVSFSDAADKIFTGGVNDVKVMDLR-----GEA 214
QY 269 TRKL-GYSSLI-----LDSTGSTLFPANCTDDNIYFMNMT--GLKTSVPAIFNGHONS--TF 319
DB 215 TMTLEGHQDTITGMSLSPDGSYLLTNGMDNKLVCVMDMRPYAPQNRCKVIFEGHQHNFKN 274
QY 320 YVKSSLSPPDQFLVSGSSDEAAYIKVS---TPWQPTVLLGHSEVTSVCWCPSDFTKI 376
DB 275 LLKCSWSPDGTKVYTAGSSDRWVHIWDTTSRTIYKLP-----GHTGSVNECVFHPTE-PII 329
QY 377 ATCSDNTL 385
DB 330 GSCSSDKNI 338

RESULT 2
US-09-590-540-5
; Sequence 5, Application US/09590540
; Patent No. 6410267
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; Lal, Preeti
; Guegler, Karl J.
; Corley, Neil C.
; Patterson, Chandra
; TITLE OF INVENTION: HUMAN GTPASE-ASSOCIATED PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Genomics, Inc.
; STREET: 3160 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 3.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/590,540
; FILING DATE: 08-Jun-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/063,743
; FILING DATE: April 21, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: David G. Streeter
; REGISTRATION NUMBER: 43,168
; REFERENCE/DOCKET NUMBER: PF-0508-1 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 343 amino acids
; TYPE: amino acid
;
```

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 2281093
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-590-540-5

Query Match 6.6%; Score 256.5; DB 2; Length 343;
Best Local Similarity 25.6%; Pred. No. 1.9e-12;
Matches 79; Conservative 44; Mismatches 117; Indels 69; Gaps 9;

QY 89 KCFKEMAHNAVFDLAWVPGELKLVTAAGDQTAKEFMDVKAGELIGTCKGHQCSLKSVAF 148
DB 87 KFMVLKGHNAILDLHWTSDGSQIVSASPKTKVRAWDVETGKQIKMAEHSFVNSCCP 146
QY 149 SKFEKAVFCTGGRGNIMVWDTRCNKKDGFYRQVNIQISGAHNTSDKQTPSKPKKQNSKG 208
DB 147 TRRGPPLIISGDDGTAKLWDMR-----QRG 172
QY 209 LAPSVDFQSVTVVLVFDQENTLVLSAGAVDGIKWDLRKNYTAHQEPIASKSFYLPSS 268
DB 173 AIQTPDKYQITAVSFSDAADKIFTGGVNDVKVMDLR-----GEA 214
QY 269 TRKL-GYSSLI-----LDSTGSTLFPANCTDDNIYFMNMT--GLKTSVPAIFNGHONS--TF 319
DB 215 TMTLEGHQDTITGMSLSPDGSYLLTNGMDNKLVCVMDMRPYAPQNRCKVIFEGHQHNFKN 274
QY 320 YVKSSLSPPDQFLVSGSSDEAAYIKVS---TPWQPTVLLGHSEVTSVCWCPSDFTKI 376
DB 275 LLKCSWSPDGTKVYTAGSSDRWVHIWDTTSRTIYKLP-----GHTGSVNECVFHPTE-PII 329
QY 377 ATCSDNTL 385
DB 330 GSCSSDKNI 338

RESULT 3
US-09-063-743-1
; Sequence 1, Application US/09063743
; Patent No. 6242214
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Patterson, Chandra
; TITLE OF INVENTION: HUMAN GTPASE-ASSOCIATED PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/063,743
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Carrone, Michael C
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0508 US
; TELECOMMUNICATION INFORMATION:
;
```

TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LUNGN0T01
CLONE: 059953
US-09-063-743-1

Query Match 6.5%; Score 253; DB 2; Length 375;
Best Local Similarity 25.7%; Pred. No. 4.3e-12;
Matches 98; Conservative 49; Mismatches 159; Indels 76; Gaps 14;

QY 20 SSQYPLQSLLTGYQCSGNDEHTSYGETGVPPFGCTFSSAPNMEHVLAVANEEGFVRLY 79
DB 54 SLQAPIM-LLSGHE-----GEV-----YCKKFH--PN-GSTLASAGFDRLLILW 93

QY 80 NTESQSFRRKCKEWAHNAVFDLAWVPELKLVTAGDQTAKEFVWVKAGELIGTCKGH 139
DB 94 NVYGCDCNYATLK---GHSGAVMELHYNTDGSMLFSASTDKTVAVWVWSETGERVKRLKGH 150

QY 140 QCSLKSVAPEKFAVCTGGRDNIMVWDTTCNKKDGFYRQVNOISGAHNTSDKQTPSK 199
DB 151 TSFVNSCYPARRGPQLVCTGSDGDKLWDIR--KKA-----IQTQNT----- 193

QY 200 PKKQNSKGLAPVDVQQSVTVVLFQDENTLVSAAGVDGIKVDLRKNVYATYRQPIAS 259
DB 194 -----YQVLAVTFNDTSDQIISGGIDNDIKVDLRQNKLYTMRGHAD 236

QY 260 KSFLYPGSSTRKLGYSLLDSTGLFANCTDDNIYMFNMTGL--KTSFVAIFNG--HQ 315
DB 237 -----SVTGLSLSSGYSLLSNAMNTVRVWVDFPAPKRCVKIFQGNVHN 283

QY 316 NSTFYVKSLSLPPDQFLVSGSSDEAAYIKWVSTPWPQPTVLLGHSEVTSVCWCPSDFTK 375
DB 284 FEKNLRCWSVDPGSKIAAGSADRFVYVWD--TTSRRILYKLPGHAGSINEVAFHP--DEPI 341

QY 376 IATCSDNTLKIWLNRGLEEK 397
DB 342 IISASSDKELYMGFESEDMDWK 363

RESULT 4
US-09-590-540-1
Sequence 1, Application US/09590540
Patent No. 6410267
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
Lal, Preeti
Guegler, Karl J.
Corley, Neil C.
Patterson, Chandra
TITLE OF INVENTION: HUMAN GTPASE-ASSOCIATED PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Genomics, Inc.
STREET: 3160 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 3.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/590,540

FILING DATE: 08-Jun-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/063,743
FILING DATE: April 21, 1998
ATTORNEY/AGENT INFORMATION:
NAME: David G. Streeter
REGISTRATION NUMBER: 43,168
REFERENCE/DOCKET NUMBER: PF-0508-1 DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 375 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LUNGN0T01
CLONE: 059953
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-590-540-1

Query Match 6.5%; Score 253; DB 2; Length 375;
Best Local Similarity 25.7%; Pred. No. 4.3e-12;
Matches 98; Conservative 49; Mismatches 159; Indels 76; Gaps 14;

QY 20 SSQYPLQSLLTGYQCSGNDEHTSYGETGVPPFGCTFSSAPNMEHVLAVANEEGFVRLY 79
DB 54 SLQAPIM-LLSGHE-----GEV-----YCKKFH--PN-GSTLASAGFDRLLILW 93

QY 80 NTESQSFRRKCKEWAHNAVFDLAWVPELKLVTAGDQTAKEFVWVKAGELIGTCKGH 139
DB 94 NVYGCDCNYATLK---GHSGAVMELHYNTDGSMLFSASTDKTVAVWVWSETGERVKRLKGH 150

QY 140 QCSLKSVAPEKFAVCTGGRDNIMVWDTTCNKKDGFYRQVNOISGAHNTSDKQTPSK 199
DB 151 TSFVNSCYPARRGPQLVCTGSDGDKLWDIR--KKA-----IQTQNT----- 193

QY 200 PKKQNSKGLAPVDVQQSVTVVLFQDENTLVSAAGVDGIKVDLRKNVYATYRQPIAS 259
DB 194 -----YQVLAVTFNDTSDQIISGGIDNDIKVDLRQNKLYTMRGHAD 236

QY 260 KSFLYPGSSTRKLGYSLLDSTGLFANCTDDNIYMFNMTGL--KTSFVAIFNG--HQ 315
DB 237 -----SVTGLSLSSGYSLLSNAMNTVRVWVDFPAPKRCVKIFQGNVHN 283

QY 316 NSTFYVKSLSLPPDQFLVSGSSDEAAYIKWVSTPWPQPTVLLGHSEVTSVCWCPSDFTK 375
DB 284 FEKNLRCWSVDPGSKIAAGSADRFVYVWD--TTSRRILYKLPGHAGSINEVAFHP--DEPI 341

QY 376 IATCSDNTLKIWLNRGLEEK 397
DB 342 IISASSDKELYMGFESEDMDWK 363

RESULT 5
US-08-190-802A-66
Sequence 66, Application US/08190802A
Patent No. 5519003
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
APPLICANT: Ron, Dorit
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dehlinger & Associates
STREET: P.O. Box 60850
CITY: Palo Alto
STATE: CA

```
/ COUNTRY: USA
/ ZIP: 94306-0850
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ FILING DATE: 01-FEB-1994
/ CLASSIFICATION: 530
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fabian, Gary R.
/ REGISTRATION NUMBER: 33,875
/ REFERENCE/DOCKET NUMBER: 8600-0139
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 324-0880
/ TELEFAX: (415) 324-0960
/ INFORMATION FOR SEQ ID NO: 66:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 514 amino acids
/ TYPE: amino acid
/ TOPOLOGY: unknown
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ INDIVIDUAL ISOLATE: YCW2 PROTEIN, Fig. 49
/ US-08-190-802A-66

Query Match
Best Local Similarity 24.1%; Score 245.5; DB 1; Length 514;
Matches 116; Conservative 62; Mismatches 196; Indels 107; Gaps 23;

QY 13 GVLR-NGWSSQYPLQSLTGYQCSGNDEHTSYGETGVVP-PFGCTFSSAPNMEHVLAV- 69
DB 47 GALTVPGAISEKQLELLN--QLNGTSDD-----PVPYTFSCITQKKASDPVKITD 96
QY 70 -----ANEGFVRLYNTESSQSPRKKCFKEWMA---HWNAPVFLAWVP-GE 110
DB 97 ITDNLVSSLLKPGYNSTEDQITLTPRAVFKVPVTRSSSAIAGHGSTILCSAFAPHTS 156
QY 111 LKLVTAAGDQTAQFWDVVRAGELIGTCKGHCQSLKSVAFSKPEKAVFCTGGRDGNIMVWDT 170
DB 157 SRMVTGAGDNTARINWCDTQTPMHTLKGHYNWVLCVSWSP-DGEVIATGSDMTIRLWDP 215
QY 171 RCKKKDGFYRQVNOISGAHN---TSDKQTP-----SKPKKKQNSKGLAPSV----- 213
DB 216 KSGQCLG-----DALRG-HSKWITSLSWEPHLVKGPKPRLASSSSKDTIKIWDTVSRV 269
QY 214 -----DFQSQSVTVVLFQDENTLVSAAGVDGIKVKWDLR-----KNYTAIRQEP 258
DB 270 COYTMSGHTNSVSVKMGQGLLYS-GSHDRTRVRVWDINSQGRCLNLIKSHAHWNHLSL 328
QY 259 SKSFLY-----PGSSTRKL---GYSSLILDSTGTSTLFANCTDDNNIYFNMTG 302
DB 329 STDYALRIGAFDHTGCKPSTPEEAQKALENYEKICKKNGNSEEMVMTASDDYTWFLWNP 388
QY 303 LK-TSPVATFNHONSTFVYKSLSPDDQFLVSGSSDEAAIYWK-----VSTPWQPPTV 355
DB 389 LKSTRPIARMTGHOKLVNHV--AFSPDGRYIYVASFDSINIKLWDGRDGRGKFIST----- 439
QY 356 LLGHSEQVTSVCMPCSPDFTKIATCSDDNTLKIW--RLNRGLBEKPGDGLKSLTVGMAQKK 413
DB 440 FRGHIASTVQVAV--SSDCRLLVSCSKDTTLKVDVTRKLSVDLPGLIKTKLYVDNSVDGK 498
QY 414 K 414
DB 499 R 499

RESULT 6
US-08-477-346-66
```

```
/ Sequence 66, Application US/08477346
/ Patent No. 6262023
/ GENERAL INFORMATION:
/ APPLICANT: Mochly-Rosen, Daria
/ APPLICANT: Ron, Dorit
/ TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
/ TITLE OF INVENTION: Thereof
/ NUMBER OF SEQUENCES: 265
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Morrison & Foerster
/ STREET: 2000 Pennsylvania Avenue, NW
/ CITY: Washington
/ STATE: DC
/ COUNTRY: USA
/ ZIP: 20006-1812
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/477,346
/ FILING DATE: 07-JUN-1995
/ CLASSIFICATION: 514
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/487,072
/ FILING DATE: 07-JUN-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: MURASHIGE, KATE H.
/ REGISTRATION NUMBER: 29,959
/ REFERENCE/DOCKET NUMBER: 2550-0025.20
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 887-1500
/ TELEFAX: (202) 887-0763
/ INFORMATION FOR SEQ ID NO: 66:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 514 amino acids
/ TYPE: amino acid
/ TOPOLOGY: unknown
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ INDIVIDUAL ISOLATE: YCW2 PROTEIN, Fig. 49
/ US-08-477-346-66

Query Match
Best Local Similarity 24.1%; Score 245.5; DB 2; Length 514;
Matches 116; Conservative 62; Mismatches 196; Indels 107; Gaps 23;

QY 13 GVLR-NGWSSQYPLQSLTGYQCSGNDEHTSYGETGVVP-PFGCTFSSAPNMEHVLAV- 69
DB 47 GALTVPGAISEKQLELLN--QLNGTSDD-----PVPYTFSCITQKKASDPVKITD 96
QY 70 -----ANEGFVRLYNTESSQSPRKKCFKEWMA---HWNAPVFLAWVP-GE 110
DB 97 ITDNLVSSLLKPGYNSTEDQITLTPRAVFKVPVTRSSSAIAGHGSTILCSAFAPHTS 156
QY 111 LKLVTAAGDQTAQFWDVVRAGELIGTCKGHCQSLKSVAFSKPEKAVFCTGGRDGNIMVWDT 170
DB 157 SRMVTGAGDNTARINWCDTQTPMHTLKGHYNWVLCVSWSP-DGEVIATGSDMTIRLWDP 215
QY 171 RCKKKDGFYRQVNOISGAHN---TSDKQTP-----SKPKKKQNSKGLAPSV----- 213
DB 216 KSGQCLG-----DALRG-HSKWITSLSWEPHLVKGPKPRLASSSSKDTIKIWDTVSRV 269
QY 214 -----DFQSQSVTVVLFQDENTLVSAAGVDGIKVKWDLR-----KNYTAIRQEP 258
DB 270 COYTMSGHTNSVSVKMGQGLLYS-GSHDRTRVRVWDINSQGRCLNLIKSHAHWNHLSL 328
QY 259 SKSFLY-----PGSSTRKL---GYSSLILDSTGTSTLFANCTDDNNIYFNMTG 302
DB 329 STDYALRIGAFDHTGCKPSTPEEAQKALENYEKICKKNGNSEEMVMTASDDYTWFLWNP 388
```



```

Best Local Similarity 24.1%; Pred. No. 2.8e-11;
Matches 116; Conservative 62; Mismatches 196; Indels 107; Gaps 23;

Qy 13 GVLR-NGWSSQYPLQSLLTGYQSGNDHSTSYGETGVVP-PFGCTTFSSAPNMEHVLAV- 69
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 47 GALRVPGAISKEQLEELLN--QLNGTSD-----PVPYTFSTIOGKKASDPVKTD 96
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 70 -----ANEEGFVRLYNTESQSRKKCFKEMMA-----HNAVFDLAWVP-GE 110
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 97 ITDNLYSGLIKPGYNSTBDQITLLYTPRAVFKVPVTRSSAIAAGHSGTILCSAFAPHTS 156
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 111 LKLVTRAAGDQTAKEWDVKAGELIGTCCKGHCQSLKSAFSEKFEKAVFCTGGRDGNIMWDT 170
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 157 SRMVTGAGDNTARIWDCDTQPMHTLKGHYNNWLCVSWSP-DGEVIATGSMNDTIRLWDP 215
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 171 RCNKKDGFYRQVNOISGAHN--TSDKQTP-----SKPKKKONSKGLAPSV----- 213
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 216 KSGQCLG-----DALRG-HSKWITSLSEPHILVKPGSKPLRASSKDGTTIKWDVTSRV 269
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 214 -----DFQSQVTVVLFODENTLVSAGAVDGIKIVDRL-----KNVTAYRQBPFA 258
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 270 QYTMISGHTNSVSCVKMGQGLLYS-GSHDRTVRWIDINSQGRICINILKSHAHWNVHLSL 328
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 259 SKSFY-----PGSSTRKL--GYSSLILDSTGTLFANCTDDNIYFNMWG 302
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 329 STDVALRIGAFDHTGKPKSTPEEAQKALENYEKICKNGNSEEMVVTASDQTYTMFLWNP 388
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 303 LK-TSPVAIFNGHONSTFYVVKSSLSPPDQPLVSGSSDEAAVIWK-----VSTPWQPPV 355
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 389 LKSTKPIARMTGHQKLVNHRV--AFSPDGRYIVSASFDSIKLMDGRDGKFIST----- 439
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 356 LLGHSQEVTVCWCPSPDTKIATCSDDNTLKIW--RLNRGLEEXPGGDKLSTVGWASQKK 413
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 440 FRGHIASYVQAV--SSDCELLVSCSKDTTLKVDVTRTKLSVDLPGIKTKLYVDWSVDGK 498
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 414 K 414
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 499 R 499

RESULT 9
US-09-291-170A-13
; Sequence 13, Application US/09291170A
; Patent No. 6410687
; GENERAL INFORMATION:
; APPLICANT: Vale, Ronald D.
; APPLICANT: Hartman, James J.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Assays for the Detection of Microtubule
; TITLE OF INVENTION: Depolymerization Inhibitors
; FILE REFERENCE: 18557B-000510US
; CURRENT APPLICATION NUMBER: US/09/291,170A
; CURRENT FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 60/081,734
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Thermomonaspora curvata
; FEATURE:
; OTHER INFORMATION: putative serine/threonine kinase Pkwa WD40 repeat
; OTHER INFORMATION: region
US-09-291-170A-13

Query Match 6.2%; Score 241; DB 2; Length 251;
Best Local Similarity 23.7%; Pred. No. 2.2e-11;
Matches 85; Conservative 55; Mismatches 103; Indels 116; Gaps 14;

Qy 35 SCNDHSTSYGET-----GVVPFPFGCTFSSAPNMEHVLAVANEEGFVRLYNTESQSRKKC 90
Db 2 SGDELHTTLEIGHTDWVRVAFSPDGA-----LLASGSDDATVRLWDVAAAEER---- 48
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 193 ALLAGSDDRTIRLMDVAAO-EEHTTLEGTPEVHSAFHP-EGTTLASEDGTIRIW 249

RESULT 11

```

US-09-724-592-13
; Sequence 13, Application US/09724592
; Patent No. 6699969
; GENERAL INFORMATION:
; APPLICANT: Vale, Ronald
; APPLICANT: Hartman, James
; TITLE OF INVENTION: Assays for the Detection of Microtubule Depolymerization Inhibitors
; FILE REFERENCE: UCSD-04884
; CURRENT APPLICATION NUMBER: US/09/724,592
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: PCT/US99/08086
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 60/081,734
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Thermomonospora curvata
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: putative serine/threonine kinase Pkwa WD40 repeat region
US-09-724-592-13

```

Query Match	6.2%	Score 241;	DB 2;	Length 251;
Best Local Similarity	23.7%	Pred. No. 2.2e-11;		
Matches 103		55; Mismatches 103;		
Matches 85; Conservative		Indels 116;		Gaps 14;

Qy	35	SGNDEHTSYGET-----GVPVPFPGCTGFSSAPNMEHVLA VANE EBGFVRLYNTESQSFRKCC	90
Db	2	SGDELHTEGTDWVRVAFSPDGA-----LLASGSDDATVRLWDVAAAEER----	48
Qy	91	FKEXMAHNAVFDLAWVPSELKLVTAAGDOTAKFWDVKAGELIGTCKGHQCSLKSVAFSK	150
Db	49	-AVPEGHYHVLIDIAFSPDGSVMVASGSRDGTARLNNVATGTEHAVLKGHTDYVYVAFSP	107
Qy	151	FEKAVFCTGRGDNIMWDTNRCKNDGFFRVQVNOISGAHNTSDQTPSKPKKKQNSKGLA	210
Db	108	DGSMVASGSRDGTIRLWDVATGKE-----RDVLQ-----A	137
Qy	211	PSVDFOQSPTVVLFDQENTLVLSAGAVDGIKIKWDLRKNTAYRQBPFIASKSFLYPGSSTR	270
Db	138	PA-----ENVVSLAFSPDGSMLVHGS-DSTVHLWD-----VASGEALH-----	176
Qy	271	KLGYSSLILDSTGSTLFLANCTDDNIYFMNMTGLKTSPIAIFNGHQNSTFYVYKS-SLSPPD	329
Db	175	-----TPEGH-----TDWVRVAFSPDG	197
Qy	330	QFLVSGSDEAAIYKWKPTFPQPTVVLGHQSQEVTSVCWCPSPDFTFKIATCSDDNLTAKI	388
Db	193	ALLASGSDRTIRLMDVAAQ-EHEHTLEGHTEPHVSHVAFHP-EGTTLASAGEDGTIRI	249

RESULT 12

US-09-673-222-13
 ; Sequence 13, Application US/09673222
 ; Patent No. 6872537
 ; GENERAL INFORMATION:
 ; APPLICANT: Vale, Ronald
 ; APPLICANT: Hartman, James
 ; TITLE OF INVENTION: Assays for the Detection of Microtubule Depolymerization Inhibitors
 ; FILE REFERENCE: UCSD-04765
 ; CURRENT APPLICATION NUMBER: US/09/673,222
 ; CURRENT FILING DATE: 2000-10-13
 ; PRIOR APPLICATION NUMBER: PCT/US99/08086
 ; PRIOR FILING DATE: 1999-04-13
 ; PRIOR APPLICATION NUMBER: 60/081,734

```

; PRIOR FILING DATE: 1998-04-14
;
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Thermomonospora curvata
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: putative serine/threonine kinase Pkwa WD40 repeat region
US-09-673-222-13

```

Query Match	6.2%;	Score 241;	DB 2;	Length 251;
Best Local Similarity	23.7%;	Pred. No. 2.2e-11;		
Matches	85;	Conservative 55;	Mismatches 103;	Indels 116; Gaps 14;
Qy	35	SGNDEHTSGET---	GVPPPPGCTFSSAPNMEHVLAVANEGFVRLYNTESQSPKKC	90
Db	2	SGDELHTEGHTDWMVRAVAFSPDGA	-----LLASGDDATVRLMDVAAEER---	48
Qy	91	FKEWMAHNAVFDLAWVPGBELKLVTAAGDOTAKFWDVKAGELIGCTCKGHOCSLKSAFVK	150	
Db	49	-AVFEQTHYVLIDIAFSPDGSVMASGSRDGTARLNNVATGTEHAVLKGHTDYVVAFAFSP	107	
Qy	151	FEKAVFCTGRGDGNIMWTRCNKKDGFYQVQNOISGAHNTSDKQTFSPKPKQNSKGLA	210	
Db	108	-DGSVMASGSRDGTIRLMDVATGKE---	RDVLQ-----A	137
Qy	211	PSYDFQGSVTVLFDQENTLVSAGVDGIIFKVDLRKNYTYARQEPITASKSFYPGSSSTR	270	
Db	138	PA-----ENVVSLAFSPDGSMLVHGS-DSTVHLMD-----	VASGEALH-----	174
Qy	271	KLGYSSILILDSTGTLFANCTDDNIYFMNMTGLKTSPIAFNGHQNSTFYVKS-SLSPPDD	329	
Db	175	-----TFEGH-----	TDWMVRAVAFSPDG	192
Qy	330	QFLVSGSSDEAAIYIKVSTPMQPPTVLLGHISQEVTSVCWCPSPFTKIATCSDNNTLKIW	388	
Db	193	ALLASGSDRTIRLMDVAAG-EBHHTLEGTPEPVSVAFHP-EGTTILASASEDGTIRIW	249	

RESIST, T 13

```

US-08-883-534-3
; Sequence 3, Application US/08883534
; Patent No. 5846777
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: TWO NEW WD-40 PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/883,534
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749

```

```
; REFERENCE/DOCKET NUMBER: PF-0332 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 606 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: NEUTGNT01
; CLONE: 1221143
; US-08-883-534-3

Query Match          6.0%; Score 233.5; DB 1; Length 606;
Best Local Similarity 20.9%; Pred. No. 3.4e-10;
Matches 97; Conservative 79; Mismatches 157; Indels 131; Gaps 20;

QY 35 SGNDHTSYGETGVPPFGCTFSSAPNMHV-----LAVANEFGVRLYNTES 83
Db 164 TGSDDNCAAFPEG---PPFKFTVGDHSRFFVNCVRFSPDGNRFATASADGQIYI--DG 218
QY 84 QSFRRKC-PKEWMAHNAVFDLAWVPGELKLVTAAGDQTAKEFDVKAGELIGT----- 135
Db 219 KTGEKVCALGGSKAHDDGGIYAISWSPDTHLLSASGDKTSKIWDVSVNSVSTFPMGSTV 278
QY 136 -----CKGHQCSLKSVAFSK-PEKAVFC 157
Db 279 LDQQLGCLWQKHLLSVLSGVINYLDNRNPKPLHVIKHSKSIQCLTVHKNGKSYIY 338
QY 158 TGRDGNIMVWTRCNKKGDFYRQ--VNOISGAHNTSDKQTPSKPKKQNSKGLAPSVDF 215
Db 339 SGSHDGHINWDSSETGENDSFAGKGTNQVS--RMTVD-----ESGQLISCSMD 386
QY 216 QOSVTVLFPD-----ENTLVSAG-----AVDGIHKVWDLRKNYT-----AY 252
Db 387 TVRYTSLMLRDYSGOGVVKLDVQPKCVAVPGGYAVVVCIGQIVLLKQDKQKCFSDINPGY 446
QY 253 ROEPIASKFLYPGSSTRKLG-----YSSL-----ILDSTGTLFANCTDDNI 295
Db 447 EPBVVA----VHPGGDTVAIGVDGNVRLYSLGTTLKDEGKLEAKGPTVDVAYSHDGA 502
QY 296 YMFNMTGLKTSVP-AIFNGH-QNSTFY-----VKSSLSPPDQFLVSGSSDAAIYKVS 347
Db 503 FLAVCDASKVTVTFVSADGYSENNVPYGHAKIVCLAWSPDNEHFASGGMWMMVYVWTLS 562
QY 348 TPWPPTVLLGHS-QEVTSCWCPSPDFTKIATCSDDNLTIKIWL 390
Db 563 DPETRVKIQAARLHHVSSLAWL--DEHTLVTTSHDASVKEWTI 604

RESULT 14
US-09-204-764-3
; Sequence 3, Application US/09204764
; Patent No. 6025464
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: TWO NEW WD-40 PROTEINS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
```

```
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7082
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7082

Query Match          6.0%; Score 233.5; DB 2; Length 627;
Best Local Similarity 20.9%; Pred. No. 3.6e-10;
Matches 97; Conservative 79; Mismatches 157; Indels 131; Gaps 20;

QY 35 SGNDHTSYGETGVPPPGCTTSSAPNMEHV-----LAVANEFGVRLYNTE 83
Db 185 TGSDDNCAAFPEG---PPFFKFTVGDHSRFVNCVRFSPDGNRFATASADGQIYIY--DG 239
QY 84 QSFRRKC-PKEWMAHNAVFDLAWPGELKLVTAAGDOTAKFWDVKAGELIGT----- 135
Db 240 KTEKVCALGSKAHGDIYVAISWSPDSTHLSASGDKTSKIWDVSVNSVSTFPMGSTV 299
QY 136 -----CKHQCSLKSVAFSK-FEKAVFC 157
Db 300 LDQQLGCLWQKDHLLSVLSGYINYLDRNPNKPLHVIKGHSKSIQCLTVHKNKGKSYIY 359
QY 158 TGGRDGNIMWTRCNKKGDFYRQ--VNOISGAHNTSDKOTPSKPKKQNSKGLAFSVD 215
Db 360 SGSHDGHINTWDSGTENDSFAGKGTNQVS--RMTVD-----ESGQLISCSMD 407
QY 216 QQSIVTVLVFOD-----ENTLVSA-----AVDGLIKYWDLRKNYT---AY 252
Db 408 TVRYTSLMLRDYSGQGVVVKLDVPKCVAVGPGGYAVVVCIGQIVLLKDKRRCFSINDPGY 467
QY 253 RQEPFASKSPYPGSSSTRKLG-----YSSL-----ILDSTGSTLFANCTDDNI 295
Db 468 EPEVVA---VHFGDITVAIGGVDGNVRLYSILGTTLKDSGKLEAKGPVTDVAYSHDGA 523
QY 296 YMFNMTGLKTSVP-AIFNGH-QNSTFY-----VKSLSPPDQFLVSGSSDEAAIYWKVS 347
Db 524 FLAVCDASKVTVTVFSVADGYSENNVFYGHAKIVCLAWSPDNEHFASGGMDMMVYVVTLS 583
QY 348 TPWQPTVLLGHS-QEVTSVCMCPSPDTKIATCDDNTLKIWL 390
Db 584 DPETRVKIQDAHLHHVSSLAWL--DEHTLVTTSHDASVKEWTI 625
```

Search completed: November 8, 2006, 18:15:39
Job time : 55 secs

This Page Blank (uspto)

GenCore version 5.1.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: November 8, 2006, 18:04:36 ; Search time 200 Seconds
(without alignments)
1668.840 Million cell updates/sec

Title: US-10-726-160-2

Perfect score: 3877

Sequence: 1 MLFNSVLRQPLQGLVLRNGWS.....TYFHRKSOEDFCGPEHSTEL 730

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

A_Geneseq.8.*

- 1: geneseqp1980s.*
- 2: geneseqp1990s.*
- 3: geneseqp2000s.*
- 4: geneseqp2001s.*
- 5: geneseqp2002s.*
- 6: geneseqp2003as.*
- 7: geneseqp2003bs.*
- 8: geneseqp2004s.*
- 9: geneseqp2005s.*
- 10: geneseqp2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3877	100.0	730	6	ABU09611 Human ret
2	3877	100.0	730	7	Adf69740 Human ret
3	3877	100.0	730	8	Ado20069 Human PRO
4	3877	100.0	730	8	Ado20232 Human PRO
5	3877	100.0	730	9	Adx69661 Human ret
6	3877	100.0	730	9	Ady17438 PRO poly
7	3877	100.0	730	10	Aee32574 Human L2D
8	3873	99.9	730	4	Aam25224 Human pro
9	3503.5	90.4	673	8	Adq97697 Human can
10	3479.5	89.7	729	6	Abu09610 Mouse can
11	2913.5	75.1	786	8	Adq97694 Mouse can
12	2293	59.1	434	4	AAE92779 Human pro
13	1439	37.1	332	4	Adm20092 Protein e
14	1413	36.4	287	4	Admi9846 Protein e
15	1216	31.4	232	4	AAE92715 Human pro
16	1165	30.0	522	8	Ado57307 Kidney de
17	1084	28.0	206	4	ABE50173 Human tra
18	1020	26.3	211	4	ABE95316 Human pro
19	793.5	20.5	769	4	ABE5620 Drosophill
20	759	19.6	146	8	Adq97699 Human can
21	587	15.1	523	8	Adt58375 Plant pol
22	406	10.5	111	4	Aag73994 Human col
23	312	8.0	610	8	AdS41755 Bacterial

24	298	7.7	1005	8	ADS41812	Ads41812 Bacterial
25	295.5	7.6	579	8	ADS29860	Ads29860 Bacterial
26	279.5	7.2	1118	8	ADS29812	Ads29812 Bacterial
27	269	6.9	608	8	ADS41794	Ads41794 Bacterial
28	267	6.9	1155	8	ADS30836	Ads30836 Bacterial
29	266	6.9	606	8	ADS29874	Ads29874 Bacterial
30	264.5	6.8	580	8	ADR86389	Adr86389 Aspergill
31	261.5	6.7	357	6	ABO53002	Abo53002 Human spl
32	261	6.7	478	8	ADS29872	Ads29872 Bacterial
33	259.5	6.7	357	4	AAm93201	Aam93201 Human pol
34	259.5	6.7	357	8	ADL30556	Adl30556 Human pro
35	258.5	6.7	1117	8	ADS29932	Ads29932 Bacterial
36	256	6.6	348	9	ABE27318	Abe27318 Pinus rad
37	256	6.6	515	6	ABR53774	Abt53774 Protein s
38	256	6.6	515	7	ADK63846	Adk63846 Disease t
39	256	6.6	515	8	ADN18878	Adn18878 Bacterial
40	255.5	6.6	655	8	ADS41636	Ads41636 Bacterial
41	254.5	6.6	1136	8	ADS30013	Ads30013 Bacterial
42	254	6.6	379	8	ADN25500	Adn25500 Bacterial
43	253	6.5	375	4	AAE02854	Aae02854 Human GTP
44	253	6.5	375	5	AAE26083	Aae26083 Human GTP
45	253	6.5	375	6	ABG75809	Abg75809 Human GTP

ALIGNMENTS

RESULT 1

ABU09611
ID ABU09611 standard; protein; 730 AA.

XX

AC ABU09611;

XX

DT 01-JUL-2003 (first entry)

XX

DE Human retinoic acid-regulated nuclear matrix-associated protein.

XX

KW Nucleic acid identification; T cell receptor; T cell; tumour antigen;

KW antigen presenting cell; HLA; nucleic acid library; ELISPOT assay;

KW enzyme-linked immunospot; T cell receptor binding epitope;

KW cancer associated antigen; ramp; human; polytope;

KW retinoic acid-regulated nuclear matrix-associated protein.

XX

OS Homo sapiens.

XX

PN US2003003485-A1.

XX

PD 02-JAN-2003.

XX

PF 14-MAY-2002; 2002US-00145396.

XX

PR 15-MAY-2001; 2001US-0291125P.

XX

PA (LUDW-) LUDWIG INST CANCER RES.

XX

PI Uenaka A, Nakayama E;

XX

WPI; 2003-416593/39.

DR

N-PSDB; ABX95498.

XX

Identifying nucleic acids encoding epitopes that bind T cell receptors on

T cells when presented by human leukocyte antigen by using an ELISPOT

assay that detects factors secreted by T cell in response to the receptor

PT binding.

XX

Example 6; Fig 13; Sipp; English.

XX

The invention describes a method of identifying nucleic acid encoding an

epitope that binds to a T cell receptor on T cell. The method involves

co-culturing antigen presenting cells that contain a nucleic acid library

and express HLA molecules presenting epitopes, with a T cell having a T

cell receptor that binds the epitope so that the T cell receptor binds

the epitope encoded by the nucleic acid library, detecting a factor

CC secreted by the T cell in response to T cell receptor binding, using
CC ELISPOT (enzyme-linked immunospot) assay, and correlating factor
CC secretion with presence of a nucleic acid encoding the epitope. The
CC method is useful for identifying a nucleic acid molecule encoding the
CC epitope that specifically binds to a T cell receptor on a T cell when
CC presented by an HLA molecule. A second method described in the invention
CC is useful for identifying antigens that specifically bind to a T cell
CC receptor on a T cell. The method is useful for identifying cancer
CC associated antigens. This is the amino acid sequence of the human
CC retinoic acid-regulated nuclear matrix-associated protein (ramp) peptides
CC of which can be combined with peptides of one or more other cancer-
CC associated antigens to form polytopes
XX

SQ Sequence 730 AA;

Query Match	100.0%;	Score 3877;	DB 6;	Length 730;
Best Local Similarity	100.0%;	Pred. No. 1.4e-293;		
Matches 730;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MLFNSVLRQPOLGVLRNGWSSQYPLQSLLTGYQCSGNDHTSYGETGVPPPPGCTFSSA	60	
Db	1	MLFNSVLRQPOLGVLRNGWSSQYPLQSLLTGYQCSGNDHTSYGETGVPPPPGCTFSSA	60	
QY	61	PNMEHVLAVANEEGFVRLYNTESQSFRRKCKFEKMAHNAVFDLAWVPGELKLVTAAGDQ	120	
Db	61	PNMEHVLAVANEEGFVRLYNTESQSFRRKCKFEKMAHNAVFDLAWVPGELKLVTAAGDQ	120	
QY	121	TAKFWDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIMVMDTRCNKXKDGFR	180	
Db	121	TAKFWDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIMVMDTRCNKXKDGFR	180	
QY	181	QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQOQSVTVVLFQDENTLVSAGAVDGI	240	
Db	181	QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQOQSVTVVLFQDENTLVSAGAVDGI	240	
QY	241	KWDLRKNTAYROEPIASKSFLYPGSSSTRKLGYSLLDSTGSTLFPANCTDDNIIYFNM	300	
Db	241	KWDLRKNTAYROEPIASKSFLYPGSSSTRKLGYSLLDSTGSTLFPANCTDDNIIYFNM	300	
QY	301	TGLKTSFPAIFNGHQNSTFYVKSLLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS	360	
Db	301	TGLKTSFPAIFNGHQNSTFYVKSLLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS	360	
QY	361	QEVTSVWCPSDFTKIATCSDNTLKIWLRLNRLGEEKPGDGLSTVGVASQKKESRPGI	420	
Db	361	QEVTSVWCPSDFTKIATCSDNTLKIWLRLNRLGEEKPGDGLSTVGVASQKKESRPGI	420	
QY	421	VVTSSQSTPAKAPRVKCNPSNSSPSSAACAPSCAGDLPLPSNTPTFSIKTSPAKARSP	480	
Db	421	VVTSSQSTPAKAPRVKCNPSNSSPSSAACAPSCAGDLPLPSNTPTFSIKTSPAKARSP	480	
QY	481	NRGVSYSVPKPPSPFKMSIRNWVTRTPSSSPITPPASETKIMSPRKALIPVSKQSQ	540	
Db	481	NRGVSYSVPKPPSPFKMSIRNWVTRTPSSSPITPPASETKIMSPRKALIPVSKQSQ	540	
QY	541	ABACSSRNVRKRRDLSSCLSVKQKCVKSCNCVTLEDQOVENLHLDLCLAGNQBDSL	600	
Db	541	ABACSSRNVRKRRDLSSCLSVKQKCVKSCNCVTLEDQOVENLHLDLCLAGNQBDSL	600	
QY	601	DSLGPTKSKIKCAGTSTISEPPSPISPYASESGTLPPLRPGCGSEGMVGKENS PENK	660	
Db	601	DSLGPTKSKIKCAGTSTISEPPSPISPYASESGTLPPLRPGCGSEGMVGKENS PENK	660	
QY	661	NWLLAAAKRAENPSPSPSSQTPNSRQSGKTLPSPTVITPSSMRKI CTYFHRKSQSD	720	
Db	661	NWLLAAAKRAENPSPSPSSQTPNSRQSGKTLPSPTVITPSSMRKI CTYFHRKSQSD	720	
QY	721	FCGPEHSTEL 730		
Db	721	FCGPEHSTEL 730		

AD69740
ID AD69740 standard; protein; 730 AA.
XX
AC AD69740;
XX
DT 12-FEB-2004 (first entry)
XX
DE Human retinoic acid (RA) regulated protein, 8.31.
XX
KW Human; retinoic acid regulated gene; retinoic acid regulated protein; RA;
KW mitogen; morphogen; stem cell progenitor cell;
KW embryonic brain development; lung; liver; kidney;
KW cancer cell differentiation; Ushers Disease; cancer; cytostatic.
XX
OS Homo sapiens.
XX
PN US2003162268-A1.
XX
PD 28-AUG-2003.
XX
PF 08-APR-2003; 2003US-00409511.
XX
PR 14-JUL-1999; 99US-00354359.
XX
PA (IPNY/) IP N Y.
PA (CHEU/) W CHEUNG W M.
XX
PI Ip NY, Cheung WMW;
XX
DR WPI; 2003-897935/82.
DR N-PSDB; AD69739.
XX
PT New retinoic acid regulated gene, useful in preparing a composition for
PT treating or preventing Ushers Disease or cancer.
XX
PS Claim 3; SEQ ID NO 2; 17pp; English.
XX
CC The present invention relates to the isolation of a novel human retinoic
CC acid (RA) regulated gene designated 8.31, and the polypeptide sequence it
CC encodes. The gene encoding the 8.31 protein maps to human chromosome
CC 1q32.1-32.2. The 8.31 protein is a mitogen or morphogen. It affects the
CC identity of stem cell progenitor cells or the development of embryonic
CC brain, lung, liver or kidney tissues. It also affects cancer cell
CC differentiation. The 8.31 protein is useful in treating or preventing
CC Ushers Disease or cancer. The present sequence represents human retinoic
CC acid (RA) regulated protein, 8.31.
XX
SQ Sequence 730 AA;

		Query Match	100.0%;	Score 3877;	DB 7;	Length 730;				
		Best Local Similarity	100.0%;	Pred. No. 1.4e-293;						
		Matches 730;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;				
QY	1	MLFNSVLRQPOLGVLRNGWSSQYPLQSLLTGYQCSGNDHTSYGETGVPPFPFCTFSSA	60							
Db	1	MLFNSVLRQPOLGVLRNGWSSQYPLQSLLTGYQCSGNDHTSYGETGVPPFPFCTFSSA	60							
QY	61	PNMEHVLAVANEEGFVRLYNTESQSFRRKCKFEKMAHNAVFDLAWVPGELKLVTAAGDQ	120							
Db	61	PNMEHVLAVANEEGFVRLYNTESQSFRRKCKFEKMAHNAVFDLAWVPGELKLVTAAGDQ	120							
QY	121	TAKFWDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIMVMDTRCNKXKDGFR	180							
Db	121	TAKFWDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIMVMDTRCNKXKDGFR	180							
QY	181	QVNQISGAHNTSDKQTPSPKPKKQNSKGLAPSVDFQOQSVTVVLFQDENTLVSAGAVDGI	240							
Db	181	QVNQISGAHNTSDKQTPSPKPKKQNSKGLAPSVDFQOQSVTVVLFQDENTLVSAGAVDGI	240							
QY	241	KWDLRKNTAYROEPIASKSFLYPGSSSTRKLGYSLLDSTGSTLIFANCTDDNIMYFNM	300							
Db	241	KWDLRKNTAYROEPIASKSFLYPGSSSTRKLGYSLLDSTGSTLIFANCTDDNIMYFNM	300							

QY 301 TGLKTSPTVAIFNGHONSTFYVKSLSPPDQFLVSGSSDEAAIWKVSTPWQPTVLLGHS 360
DB |||||
301 TGLKTSPTVAIFNGHONSTFYVKSLSPPDQFLVSGSSDEAAIWKVSTPWQPTVLLGHS 360
QY 361 QEVTSVCWCPSDFTKIATCSDNTLKIWLNRGLBEKPGDKLSTVGWASQKKESRPGL 420
DB |||||
361 QEVTSVCWCPSDFTKIATCSDNTLKIWLNRGLBEKPGDKLSTVGWASQKKESRPGL 420
QY 421 VVTSSQSTPAKAPRVKCNPSNPPSSAACAACPSACAGDPLPSNTPTFSIKTSPAKARSP 480
DB |||||
421 VVTSSQSTPAKAPRVKCNPSNPPSSAACAACPSACAGDPLPSNTPTFSIKTSPAKARSP 480
QY 481 NRRGSVSSVPKPPSSFKMSIRNWVTRTPSSPPPIPPASETKIMSPRKALIPVSKSSQ 540
DB |||||
481 NRRGSVSSVPKPPSSFKMSIRNWVTRTPSSPPPIPPASETKIMSPRKALIPVSKSSQ 540
QY 541 AEACSESRNRVRRRLDSSCLSVKQCVKSCNCVTLDQOVENLHLDLCLAGNQEDLSK 600
DB |||||
541 AEACSESRNRVRRRLDSSCLSVKQCVKSCNCVTLDQOVENLHLDLCLAGNQEDLSK 600
QY 601 DSLGPTKSKIEGAGTSISEPPSPISPYASESGTLPLRPGEGSEVMVGKENSSENK 660
DB |||||
601 DSLGPTKSKIEGAGTSISEPPSPISPYASESGTLPLRPGEGSEVMVGKENSSENK 660
QY 661 NWLLAAAKRAENPSRPSQSTNSRRQSGKTLPSPTIIPSSMRKICTYFHRKSQED 720
DB |||||
661 NWLLAAAKRAENPSRPSQSTNSRRQSGKTLPSPTIIPSSMRKICTYFHRKSQED 720
QY 721 FCGPEHSTEL 730
DB |||||
721 FCGPEHSTEL 730

RESULT 3
ADO20069
ID ADO20069 standard; protein; 730 AA.
XX
AC ADO20069;
XX
DT 12-AUG-2004 (first entry)
DE
DE Human PRO polypeptide #489.
KW Human; PRO: immune related disorder; systemic lupus erythematosus;
KW rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis;
KW systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis;
KW autoimmune haemolytic anaemia; autoimmune thrombocytopenia; thyroiditis;
KW diabetes mellitus; renal disease; demyelinating disease;
KW central nervous system; peripheral nervous system;
KW demyelinating polyneuropathy; Guillain-Barre syndrome;
KW chronic inflammatory demyelinating polyneuropathy.
XX
OS Homo sapiens.
XX
XX WO2004043361-A2.
PN
XX 27-MAY-2004.
PD
XX
PF 06-NOV-2003; 2003WO-US035268..
XX
XX 08-NOV-2002; 2002US-0425235P.
PR
XX
XX (GETH) GENENTECH INC.
PA
XX Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PW;
PI Wood WJ, Wu TD;
XX
XX WPI; 2004-420067/39.
DR
DR N-PSDB; ADO20068.
XX
PT Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
PT treating an immune related disorder such as systemic lupus erythematosus,
PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or

PT spondyloarthropathy.
XX
PS Claim 7; SEQ ID NO 978; 1731pp; English.
XX
CC The invention relates to human PRO polypeptides and the polynucleotides
CC encoding them. The polypeptides and polynucleotides are useful for
CC treating and diagnosing immune related disorders in mammals. The immune
CC related disorders include systemic lupus erythematosus, rheumatoid
CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic
CC sclerosis, Sjogren's syndrome, autoimmune thrombocytopenia, thyroiditis, diabetes
CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
CC mellitus, immune-mediated renal disease, demyelinating diseases of the
CC central or peripheral nervous system, demyelinating polyneuropathy,
CC Guillain-Barre syndrome and chronic inflammatory demyelinating
CC polyneuropathy. This sequence represents a human PRO polypeptide of the
CC invention.
XX
SQ Sequence 730 AA;
Query Match 100.0%; Score 3877; DB 8; Length 730;
Best Local Similarity 100.0%; Pred. No. 1.4e-293;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLFNSVLRLQPLGVLRNCGWSQYPLQSLLTGYQCSGNDHTSYGETGVPPVPGCTFSSA 60
DB |||||
1 MLFNSVLRLQPLGVLRNCGWSQYPLQSLLTGYQCSGNDHTSYGETGVPPVPGCTFSSA 60
QY 61 PMEHVLAVANEBSGFVRLYNTESQSFRRKCKFEMAHNAVFDLAWVPGELKLVTAAGDQ 120
DB |||||
61 PMEHVLAVANEBSGFVRLYNTESQSFRRKCKFEMAHNAVFDLAWVPGELKLVTAAGDQ 120
QY 121 TAKPMDVRAGELIGTCKGHQCSLSKSAFVKFAVCTGGRDGNIMWDTNRCNKDGFYR 180
DB |||||
121 TAKPMDVRAGELIGTCKGHQCSLSKSAFVKFAVCTGGRDGNIMWDTNRCNKDGFYR 180
QY 181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDPQQSVTVVLFQDENTLVSAGAVDGI 240
DB |||||
181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDPQQSVTVVLFQDENTLVSAGAVDGI 240
QY 241 KWDLRKNYTAHQPIASKSFYPGSSRTRKLGYSLLDSTGSLFANCTDDNLYMFNM 300
DB |||||
241 KWDLRKNYTAHQPIASKSFYPGSSRTRKLGYSLLDSTGSLFANCTDDNLYMFNM 300
QY 301 TGLKTSPTVAIFNGHONSTFYVKSLSPPDQFLVSGSSDEAAIWKVSTPWQPTVLLGHS 360
DB |||||
301 TGLKTSPTVAIFNGHONSTFYVKSLSPPDQFLVSGSSDEAAIWKVSTPWQPTVLLGHS 360
QY 361 QEVTSVCWCPSDFTKIATCSDNTLKIWLNRGLBEKPGDKLSTVGWASQKKESRPGL 420
DB |||||
361 QEVTSVCWCPSDFTKIATCSDNTLKIWLNRGLBEKPGDKLSTVGWASQKKESRPGL 420
QY 421 VVTSSQSTPAKAPRVKCNPSNPPSSAACAACPSACAGDPLPSNTPTFSIKTSPAKARSP 480
DB |||||
421 VVTSSQSTPAKAPRVKCNPSNPPSSAACAACPSACAGDPLPSNTPTFSIKTSPAKARSP 480
QY 481 NRRGSVSSVPKPPSSFKMSIRNWVTRTPSSPPPIPPASETKIMSPRKALIPVSKSSQ 540
DB |||||
481 NRRGSVSSVPKPPSSFKMSIRNWVTRTPSSPPPIPPASETKIMSPRKALIPVSKSSQ 540
QY 541 AEACSESRNRVRRRLDSSCLSVKQCVKSCNCVTLDQOVENLHLDLCLAGNQEDLSK 600
DB |||||
541 AEACSESRNRVRRRLDSSCLSVKQCVKSCNCVTLDQOVENLHLDLCLAGNQEDLSK 600
QY 601 DSLGPTKSKIEGAGTSISEPPSPISPYASESGTLPLRPGEGSEVMVGKENSSENK 660
DB |||||
601 DSLGPTKSKIEGAGTSISEPPSPISPYASESGTLPLRPGEGSEVMVGKENSSENK 660
QY 661 NWLLAAAKRAENPSRPSQSTNSRRQSGKTLPSPTIIPSSMRKICTYFHRKSQED 720
DB |||||
661 NWLLAAAKRAENPSRPSQSTNSRRQSGKTLPSPTIIPSSMRKICTYFHRKSQED 720
QY 721 FCGPEHSTEL 730
DB |||||
721 FCGPEHSTEL 730

Db 721 FCGPEHSTEL 730

RESULT 4

ID ADO20232 standard; protein; 730 AA.

XX ADO20232;

XX 12-AUG-2004 (first entry)

DT Human PRO polypeptide #567.

DE

XX Human; PRO; immune related disorder; systemic lupus erythematosus;

XX rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis;

KW systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis;

KW autoimmune haemolytic anaemia; autoimmune thrombocytopenia; thyroiditis;

KW diabetes mellitus; renal disease; demyelinating disease;

KW central nervous system; peripheral nervous system;

KW demyelinating polyneuropathy; Guillain-Barre syndrome;

KW chronic inflammatory demyelinating polyneuropathy.

XX

OS Homo sapiens.

XX WO2004043361-A2.

PN 27-MAY-2004.

XX

XX 06-NOV-2003; 2003WO-US035268.

PF

XX 08-NOV-2002; 2002US-0425235P.

PR

XX (GETH) GENENTECH INC.

PA

PI Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PW;

PI Wood WI, Wu TD;

PI

DR WPI; 2004-420067/39.

DR N-PSDB; ADO20231.

XX

XX Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for

PT treating an immune related disorder such as systemic lupus erythematosus,

PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or

PT spondyloarthropathy.

XX

XX Claim 7; SEQ ID NO 1134; 1731pp; English.

PS

XX The invention relates to human PRO polypeptides and the polynucleotides

CC encoding them. The polypeptides and polynucleotides are useful for

CC treating and diagnosing immune related disorders in mammals. The immune

CC related disorders include systemic lupus erythematosus, rheumatoid

CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic

CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune

CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes

CC mellitus, immune-mediated renal disease, demyelinating diseases of the

CC central or peripheral nervous system, demyelinating polyneuropathy,

CC Guillain-Barre syndrome and chronic inflammatory demyelinating

CC polyneuropathy. This sequence represents a human PRO polypeptide of the

CC invention.

XX

XX Sequence 730 AA;

SQ

Query Match 100.0%; Score 3877; DB 8; Length 730;

Best Local Similarity 100.0%; Pred. No. 1.4e-293;

Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLFNSVLROQLGVLNGLNGLSSQVPLQSLTGTQCSGNDEHTSYGETGVVPFPGCTFSSA 60

DB 1 MLFNSVLROQLGVLNGLNGLSSQVPLQSLTGTQCSGNDEHTSYGETGVVPFPGCTFSSA 60

QY 61 PNMEHLVANEGFVRLNNTSQQSKCFKEWMAHNAVFDLAWVPGELKLVTAAGDQ 120

DB 61 PNMEHLVANEGFVRLNNTSQQSKCFKEWMAHNAVFDLAWVPGELKLVTAAGDQ 120

QY 121 TAKFWDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIMVWDTFCNKKDGFYR 180

DB 121 TAKFWDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIMVWDTFCNKKDGFYR 180

QY 181 QVNOISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQOSVTVVLFDENTLVSAGAVDGI 240

DB 181 QVNOISGAHNTSDKQTPSKPKKQNSKGLAPSVDFQOSVTVVLFDENTLVSAGAVDGI 240

QY 241 KVDLRKNYTAIROEPIASKSFLYPGSSSTRKLGYSLLDSTGLFANCTDDNIYFMNM 300

DB 241 KVDLRKNYTAIROEPIASKSFLYPGSSSTRKLGYSLLDSTGLFANCTDDNIYFMNM 300

QY 301 TGLKTSVAIFNGHQNSTFYVKSSLPDDQFLVSGSSDEAAAYIKWVSPWPQPTVLLGHS 360

DB 301 TGLKTSVAIFNGHQNSTFYVKSSLPDDQFLVSGSSDEAAAYIKWVSPWPQPTVLLGHS 360

QY 361 QEVTSVCWCPDSDFTKIATCSDNTLKIWLNRGLKEKPGDKLSTVGWASOKKESRPL 420

DB 361 QEVTSVCWCPDSDFTKIATCSDNTLKIWLNRGLKEKPGDKLSTVGWASOKKESRPL 420

QY 421 VVTSSQSTPAKAPRVKCNPSNSPSSAACAPSCAGDLPLPSNTPTFSIKTSPAKARSP 480

DB 421 VVTSSQSTPAKAPRVKCNPSNSPSSAACAPSCAGDLPLPSNTPTFSIKTSPAKARSP 480

QY 481 NRRGVSVSVPKPPSSPKMSIRNWVTRTPSSSPPTTPPASETKIMS PRKALIPVSQKSSQ 540

DB 481 NRRGVSVSVPKPPSSPKMSIRNWVTRTPSSSPPTTPPASETKIMS PRKALIPVSQKSSQ 540

QY 541 AEACSESRNRKRLDSSCLSESVKQKVCVKNCTVLDGQVENLHLDLCCLAGNQEDLSK 600

DB 541 AEACSESRNRKRLDSSCLSESVKQKVCVKNCTVLDGQVENLHLDLCCLAGNQEDLSK 600

QY 601 DSLGPTKSSKIEGAGTISIEPPSPISPYASESCGTLPLPLPCGEGSEMVCKENSSPENK 660

DB 601 DSLGPTKSSKIEGAGTISIEPPSPISPYASESCGTLPLPLPCGEGSEMVCKENSSPENK 660

QY 661 NWLLAMAAKRAENPSPRSPSSQTPNSRRQSGKTLPSVTTTPSSMRKICTYFHRKSQED 720

DB 661 NWLLAMAAKRAENPSPRSPSSQTPNSRRQSGKTLPSVTTTPSSMRKICTYFHRKSQED 720

QY 721 FCGPEHSTEL 730

DB 721 FCGPEHSTEL 730

RESULT 5

ADK69661

ID ADK69661 standard; protein; 730 AA.

XX

AC ADK69661;

DT 05-MAY-2005 (first entry)

XX

DE Human retinoic acid regulated nuclear matrix protein.

XX

KW DNA purification; retinoic acid; diagnosis; hepatocellular carcinoma;

KW cytostatic; prognosis; cancer; liver tumor;

KW retinoic acid regulated nuclear matrix protein.

XX

OS Homo sapiens.

XX

PN US2005037372-A1.

XX

PD 17-FEB-2005.

XX

PF 02-DEC-2003; 2003US-00726160.

XX

PR 14-JUL-1999; 99US-00354359.

PR 08-APR-2003; 2003US-00409511.

XX

PA (UYHK-) UNIV HONG KONG SCI & TECHNOLOGY.

PI Ip NY, Cheung WW; 721 FCGPEHSTEL 730
XX WPI: 2005-172246/18. |||||
DR N-PSDB; ADX69660. 721 FCGPEHSTEL 730
XX
PT Novel isolated retinoic acid regulated gene, useful for screening and
XX determining prognosis of patient having Hepatocellular cancer.
XX
PS Claim 2; SEQ ID NO 2; 21pp; English.
XX
XX The invention relates to an isolated retinoic acid regulated gene (I)
CC having a fully defined 281 nucleotides sequence given in specification.
CC A protein (II) encoded by (I) is useful as a screening tool for
CC diagnosing Hepatocellular carcinomas and for monitoring treatment or
CC progression of Hepatocellular carcinomas. An antibody (III) to (II) is
CC useful for screening and determining the prognosis of a patient having
CC Hepatocellular cancer, where the biological samples comprise liver
CC tissues. This sequence corresponds to the retinoic acid regulated nuclear
CC matrix protein (RAMP; II).
XX
SQ Sequence 730 AA;

Query Match 100.0%; Score 3877; DB 9; Length 730;
Best Local Similarity 100.0%; Pred. No. 1.4e-293;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLFNSVLROPLQGLVLRNGWSSQYPLQSLITGYQCSGNDEHTSYGETGVPPPGCTFSSA 60
Db |||||
Qy 1 MLFNSVLROPLQGLVLRNGWSSQYPLQSLITGYQCSGNDEHTSYGETGVPPPGCTFSSA 60
Db |||||
Qy 61 PNMEHLVAVANEEGFVRLYNTESQSPFRKCKFKEWMAHNAVFDLAWVPGLKLVTAAGDQ 120
Db |||||
Qy 61 PNMEHLVAVANEEGFVRLYNTESQSPFRKCKFKEWMAHNAVFDLAWVPGLKLVTAAGDQ 120
Db |||||
Qy 121 TAKFMDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIMWDTRCNKKDGYR 180
Db |||||
Qy 121 TAKFMDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIMWDTRCNKKDGYR 180
Db |||||
Qy 181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFOQSSTVTLFQDENTLVSAAGVDGII 240
Db |||||
Qy 181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFOQSSTVTLFQDENTLVSAAGVDGII 240
Db |||||
Qy 241 KWDLRKNYTAIROEPIASKSFLYPGSSTRKGLYSSLLDSTGLFANCTDDNIMYFNM 300
Db |||||
Qy 241 KWDLRKNYTAIROEPIASKSFLYPGSSTRKGLYSSLLDSTGLFANCTDDNIMYFNM 300
Db |||||
Qy 301 TGLKTSVAIFNGHONSTFYVYKSSLSLDDQFLVSGSSDEAAYIWKVSTPMQPTVLLGHS 360
Db |||||
Qy 301 TGLKTSVAIFNGHONSTFYVYKSSLSLDDQFLVSGSSDEAAYIWKVSTPMQPTVLLGHS 360
Db |||||
Qy 361 QEVTSVCMCPDFTKIACTCDDNTLKIWLNRGLKEKPGDKLSTVGVASQKKESRPGI 420
Db |||||
Qy 361 QEVTSVCMCPDFTKIACTCDDNTLKIWLNRGLKEKPGDKLSTVGVASQKKESRPGI 420
Db |||||
Qy 421 VVTSSOSTPAKAPRVKCNPSNPSAAACAPSAGDLPLPSNTPPTSIIKTSAPAKARSPI 480
Db |||||
Qy 421 VVTSSOSTPAKAPRVKCNPSNPSAAACAPSAGDLPLPSNTPPTSIIKTSAPAKARSPI 480
Db |||||
Qy 481 NRRGSVSSVPKPPSSPKMIRNMTVTRTPSSPPITPPASSETKIMSPRKALIPVSQKSSQ 540
Db |||||
Qy 481 NRRGSVSSVPKPPSSPKMIRNMTVTRTPSSPPITPPASSETKIMSPRKALIPVSQKSSQ 540
Db |||||
Qy 541 AEACSESRNRVKRLDSSCLSVKQKVCNCVTELDGOVENLHLDCLLAGNQEDLSK 600
Db |||||
Qy 541 AEACSESRNRVKRLDSSCLSVKQKVCNCVTELDGOVENLHLDCLLAGNQEDLSK 600
Db |||||
Qy 601 DSLGPTKSSKIEGAGTSISRPSPSPISPVASESCOTLPLRPPCGEGSEMGKENSPEK 660
Db |||||
Qy 601 DSLGPTKSSKIEGAGTSISRPSPSPISPVASESCOTLPLRPPCGEGSEMGKENSPEK 660
Db |||||
Qy 661 NWLLAMAARKAENPSPSRSSQTPNSRRSGKTLPSFVTITPSSMRKICTYFHRKSQED 720
Db |||||
Qy 661 NWLLAMAARKAENPSPSRSSQTPNSRRSGKTLPSFVTITPSSMRKICTYFHRKSQED 720
Db |||||

Qy 721 FCGPEHSTEL 730
Db 721 FCGPEHSTEL 730
RESULT 6
ADY17438
ID ADY17438 standard; protein; 730 AA.
XX
AC ADY17438;
XX
XX 05-MAY-2005 (first entry)
XX
XX PRO polypeptide SEQ ID NO 3244.
XX
XX Antiinflammatory; Immune disorder; Dermatological; Immunosuppressive;
KW Antirheumatic; Antiarthritic; Osteopathic; Hemostatic; Antianemic;
KW Antithyroid; Antidiabetic; Nephrotropic; CNS-Gen.; Hepatotrophic;
KW Virucide; Gastrointestinal-Gen.; Antipsoriatic; Antiasthmatic;
KW Antiallergic; ds; gene; diagnosis.
XX
OS Homo sapiens.
XX
XX WO2005016962-A2.
XX
XX 24-FEB-2005.
XX
XX 11-AUG-2004; 2004WO-US026249.
XX
XX 11-AUG-2003; 2003US-0493546P.
XX
XX (GETH) GENENTECH INC.
XX
XX Abbas A, Clark H, Ouyang W, Williams MP, Wood WI, Wu TD;
XX WPI; 2005-182330/19.
XX
XX New nucleic acid encoding PRO polypeptide, useful for diagnosing and
PT treating an immune related disorder, e.g. systemic lupus erythematosus,
PT rheumatoid arthritis, osteoarthritis, thyroiditis, or diabetes mellitus.
XX
XX Claim 8; SEQ ID NO 3244; 158pp; English.
XX
XX The invention relates to an isolated nucleic acid encoding a PRO
CC polypeptide. The polypeptide, agonist or an antagonist, antibody,
CC composition, and method are useful for diagnosing and treating an immune
CC related disorder, e.g. systemic lupus erythematosus, rheumatoid
CC arthritis. The present sequence represents a DNA encoding a PRO
CC polypeptide.
XX
XX Sequence 730 AA;

Query Match 100.0%; Score 3877; DB 9; Length 730;
Best Local Similarity 100.0%; Pred. No. 1.4e-293;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MLFNSVLROPLQGLVLRNGWSSQYPLQSLITGYQCSGNDEHTSYGETGVPPPGCTFSSA 60
Db |||||
Qy 1 MLFNSVLROPLQGLVLRNGWSSQYPLQSLITGYQCSGNDEHTSYGETGVPPPGCTFSSA 60
Db |||||
Qy 61 PNMEHLVAVANEEGFVRLYNTESQSPFRKCKFKEWMAHNAVFDLAWVPGLKLVTAAGDQ 120
Db |||||
Qy 61 PNMEHLVAVANEEGFVRLYNTESQSPFRKCKFKEWMAHNAVFDLAWVPGLKLVTAAGDQ 120
Db |||||
Qy 121 TAKFMDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIMWDTRCNKKDGYR 180
Db |||||
Qy 121 TAKFMDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIMWDTRCNKKDGYR 180
Db |||||
Qy 181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFOQSSTVTLFQDENTLVSAAGVDGII 240
Db |||||
Qy 181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSVDFOQSSTVTLFQDENTLVSAAGVDGII 240
Db |||||

QY 241 KWDLRKNYAYROPIASKSFLYPGSTRKGLYSGSLILDSTGLFANCTDDNIYMFNM 300
DB |||||||
QY 241 KWDLRKNYAYROPIASKSFLYPGSTRKGLYSGSLILDSTGLFANCTDDNIYMFNM 300
DB |||||||
QY 301 TGLKTSPIVAIFNGHQNSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
DB |||||||
QY 301 TGLKTSPIVAIFNGHQNSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
DB |||||||
QY 361 QEVTSVCWCPDFTKIATCSDNTLKIWLNRGLBEKPGDKLSTVGWASQKKESRPGI 420
DB |||||||
QY 361 QEVTSVCWCPDFTKIATCSDNTLKIWLNRGLBEKPGDKLSTVGWASQKKESRPGI 420
DB |||||||
QY 421 VVTSSQSTPAKAPRVKCNPNSSPSSAACAPSCAGDLPLPNTPTFSIKTSPAKARSP 480
DB |||||||
QY 421 VVTSSQSTPAKAPRVKCNPNSSPSSAACAPSCAGDLPLPNTPTFSIKTSPAKARSP 480
DB |||||||
QY 481 NRRGVSVPKPPSPFMSIRNWTRTPSSPPITPPASETKIMSPKALIPVQKSSQ 540
DB |||||||
QY 481 NRRGVSVPKPPSPFMSIRNWTRTPSSPPITPPASETKIMSPKALIPVQKSSQ 540
DB |||||||
QY 541 ABACSESRNRVRRLDSSCLESVKQKCVKSCNCVTDLGQVENLHDLCLAGNEDLSK 600
DB |||||||
QY 601 DSIQPTKSSKIEGAGTSISEPSPISPYASESCGTLPLRRCGRGSEMVKGENSEPENK 660
DB |||||||
QY 601 DSIQPTKSSKIEGAGTSISEPSPISPYASESCGTLPLRRCGRGSEMVKGENSEPENK 660
DB |||||||
QY 661 NWLLAAAKRKAENPSRPSQTPNSRRQSGTKLPSPTIIPSSMRKICTYFHRKSQED 720
DB |||||||
QY 721 FCGPEHSTEL 730
DB |||||||
QY 721 FCGPEHSTEL 730
DB |||||||

RESULT 7
ID ABE32574
XX ABE32574 standard; protein; 730 AA.

AC ABE32574;
XX
DT 09-FEB-2006 (first entry)
DE Human L2DTL protein SEQ ID NO 8.
KW cell cycle; p53; cancer; neoplasm; cytotstatic; lung tumor; breast tumor;
KW ovarian tumor; ubiquitin; L2DTL.

OS Homo sapiens.
XX
PN WO2005114188-A2.
XX
PD 01-DEC-2005.

PF 27-APR-2005; 2005WO-US014615.
XX
PR 27-APR-2004; 2004US-0565707P.
XX
PA (UYVA) UNIV YALE.
XX
PI Zhang H, Banks DP, Higa LAA;
XX
XX WPI: 2006-020151/02.
DR GENBANK; AAF35182.

XX Identifying compound for modulating ubiquitination of p53 polypeptide by
PT incubating CUL4 E3 ligase complex with a p53 polypeptide and comparing
PT ubiquitination of p53 polypeptide in the presence or absence of compound.
XX
PS Disclosure; SEQ ID NO 8; 88pp; English.

CC This invention describes a novel method for identifying a compound which
CC modulates the ubiquitination of a p53 polypeptide. This method involves
CC incubating the CUL4 E3 ligase complex (comprising CUL4A, CUL4B, DBB1,
CC L2DTL, ROC1, ROC2, MDM2 and Pirh2, preferably Pirh2) with a p53
CC polypeptide for the ubiquitination of p53 and comparing the
CC ubiquitination of the p53 polypeptide measured in the step of incubating
CC to ubiquitination of a p53 polypeptide in a mixture not contacted with
CC the test compound to determine a difference in the ubiquitination of the
CC p53 substrate. The ligase complex is incubated with a p53 polypeptide in
CC the presence of one or more of ubiquitin-activating E1 enzymes, ubiquitin
CC -conjugating E2 enzymes, ATPs and ubiquitins. The ubiquitin-conjugating
CC E2 enzyme is chosen from Cdc34, UbcH1, UbcH2, UbcH3, UbcH4, UbcH5, UbcH6,
CC UbcH7, UbcH10 and L-UBC. The UbcH5 is UbcH5A, UbcH5B or UbcH5C. Methods
CC are also disclosed for a) decreasing the level of a p53 polypeptide in a
CC cell, by increasing the expression level of a CUL4-associated
CC polynucleotide (e.g. an E2 conjugation enzymes chosen from Cdc34 E2,
CC UbcH5B E2, CUL4A, CUL4B, DBB1, L2DTL, ROC1, ROC2, MDM2, Pirh2 and UbcH5C
CC E2) in the cell, increasing the level of a CUL4-associated polypeptide in
CC the cell, increasing binding between a p53 polypeptide and a CUL4 E3
CC ligase complex, increasing the ubiquitination activity of a CUL4 ligase
CC complex, increasing the binding between a CUL4 ligase complex and MDM2,
CC increasing the binding between a CUL4 ligase complex and Pirh2, or
CC increasing the activity of CUL1 E3 ligase complex and b) increasing the
CC level of a p53 polypeptide in a cell, which involves decreasing the
CC expression level of an mRNA encoding a CUL4-associated polypeptide in the
CC cell, decreasing the level of a CUL4-associated polypeptide in the cell,
CC decreasing binding between a p53 polypeptide and a CUL4 E3 ligase
CC complex, decreasing the ubiquitination activity of a CUL4 ligase complex,
CC decreasing the binding between a CUL4 ligase complex and MDM2, decreasing
CC the binding between a CUL4 ligase complex and Pirh2 or decreasing the
CC activity of a CUL1 E3 ligase complex. The methods of the invention are
CC useful for identifying modulators of ubiquitination of p53 polypeptide
CC utilized for treating cancer e.g. lung cancer, breast cancer or ovarian
CC carcinoma and enabling a high-throughput identification method for
CC modulators of ubiquitination of p53 polypeptide. This sequence represents
CC the human L2DTL protein, part of the CUL4 E3 ligase complex incubated
CC with p53 to investigate ubiquitination.

XX
SQ Sequence 730 AA;

Query Match 100.0%; Score 3877; DB 10; Length 730;
Best Local Similarity 100.0%; Pred. No. 1.4e-293;
Matches 730; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLFNSVLRLQPOLGVLNRGWSQYPLQSLLTGYQCSGNDHSTSYGTVGVPPFGCTFSSA 60
DB |||||||
QY 1 MLFNSVLRLQPOLGVLNRGWSQYPLQSLLTGYQCSGNDHSTSYGTVGVPPFGCTFSSA 60
DB |||||||
QY 61 PMEHVLAVANEEGFVRLYNTESQSFRRKKCKEWMHNAVFDLAWVPGELKLVTAAGDQ 120
DB |||||||
QY 61 PMEHVLAVANEEGFVRLYNTESQSFRRKKCKEWMHNAVFDLAWVPGELKLVTAAGDQ 120
DB |||||||
QY 121 TAKFWDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIWMDTRCNCKDGFYR 180
DB |||||||
QY 121 TAKFWDVKAGELIGTCKGHQCSLKSVAFSKFEKAVFCTGGRDGNIWMDTRCNCKDGFYR 180
DB |||||||
QY 181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSDVDFQQSVTVVLFQDENTLVAGAVDGI 240
DB |||||||
QY 181 QVNQISGAHNTSDKQTPSKPKKQNSKGLAPSDVDFQQSVTVVLFQDENTLVAGAVDGI 240
DB |||||||
QY 241 KWDLRKNYAYROPIASKSFLYPGSTRKGLYSGSLILDSTGLFANCTDDNIYMFNM 300
DB |||||||
QY 241 KWDLRKNYAYROPIASKSFLYPGSTRKGLYSGSLILDSTGLFANCTDDNIYMFNM 300
DB |||||||
QY 301 TGLKTSPIVAIFNGHQNSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
DB |||||||
QY 301 TGLKTSPIVAIFNGHQNSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
DB |||||||
QY 361 QEVTSVCWCPDFTKIATCSDNTLKIWLNRGLBEKPGDKLSTVGWASQKKESRPGI 420
DB |||||||
QY 361 QEVTSVCWCPDFTKIATCSDNTLKIWLNRGLBEKPGDKLSTVGWASQKKESRPGI 420
DB |||||||
QY 421 VVTSSQSTPAKAPRVKCNPNSSPSSAACAPSCAGDLPLPNTPTFSIKTSPAKARSP 480
DB |||||||

```
Db 421 VVTSSQSTPAKAPRVKCNPSNPSAAACAPSCAGDLPLPSNTPTFSIKTSPAKARSP 480
Qy 481 NRRGSVSVSPKPPSFKMSIRNWNVTRTPSSSPPIPPASETKIMSPKALIPVSQKSSQ 540
Db 481 NRRGSVSVSPKPPSFKMSIRNWNVTRTPSSSPPIPPASETKIMSPKALIPVSQKSSQ 540
Qy 541 AEACSESRNRVKRRLLDSSCLSVKQKVCVSCNCVTELDGQVENLHLDLCLAGNQBDSLK 600
Db 541 AEACSESRNRVKRRLLDSSCLSVKQKVCVSCNCVTELDGQVENLHLDLCLAGNQBDSLK 600
Qy 601 DSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLPLRPGEGSEVMVGKENSSENK 660
Db 601 DSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLPLRPGEGSEVMVGKENSSENK 660
Qy 661 NWLLAAAKRAENPSPRSPSSQTPNSRRQSGTKLPSPTITPSSMRKICTYFHRKSQED 720
Db 661 NWLLAAAKRAENPSPRSPSSQTPNSRRQSGTKLPSPTITPSSMRKICTYFHRKSQED 720
Qy 721 FCGPEHSTEL 730
Db 721 FCGPEHSTEL 730

RESULT 8
AAM25224
ID AAM25224 standard; protein; 730 AA.
XX
AC AAM25224;
XX
DT 04-DEC-2001 (first entry)
XX
DE Human prostate-related gene 83P5G4 encoded protein.
XX
KW 83P5G4-related protein; prostate; testis; tissue; cancer; bladder; liver;
KW tumour; kidney; brain; bone; ovary; breast; pancreas; colon; lung; serum;
KW cytostatic; gene therapy; antibody therapy; ribozyme; blood; cervix;
KW single chain monoclonal antibody; urine; uterus; rectum; stomach; human;
KW chromosome 1q31-q32.
XX
OS Homo sapiens.
XX
PN WO200159115-A2.
XX
PD 16-AUG-2001.
XX
PF 09-FEB-2001; 2001WO-US004426.
XX
PR 09-FEB-2000; 2000US-0181261P.
XX
PA (UROC-) UROGENESYS INC.
XX
PI Hubert RS, Afar DEH, Challita-Eid PM, Paris M, Levin E;
PI Mitchell SC, Jakobovits A;
XX
DR WPI; 2001-514669/56.
DR N-PSDB; AAR99153.
PT
PT An isolated 83P5G4-related protein useful as a diagnostic and/or
PT therapeutic agent in multiple cancers such as prostate, bladder and bone
PT cancer.
XX
PS Claim 12; Fig 2; 112pp; English.
XX
CC The polypeptide sequences represent the 83P5G4-related protein and
CC peptide fragments of the protein. 83P5G4 exhibits prostate specific
CC expression in normal adult tissue, but it is also aberrantly expressed in
CC many cancers including tumours of the prostate, testis, bladder, kidney,
CC brain, bone, cervix, uterus, ovary, breast, pancreas, stomach, rectum,
CC liver, colon and lung. The 83P5G4 polynucleotide, its related protein and
CC peptide fragments and specific PCR primers are therefore useful for
CC diagnosing and treating cancer. A vector comprising a polynucleotide
CC which encodes a single chain monoclonal antibody, that immunospecifically
```

```
CC binds to an 83P5G4-related protein, and a ribozyme capable of cleaving a
CC polynucleotide having the 83P5G4 coding sequence, are both useful in the
CC preparation of a composition for treating a patient with a cancer that
CC expresses 83P5G4. The sequences can be used in diagnostic methods to
CC monitor the level of 83P5G4 gene products in serum, blood, urine and
CC tissue and to thereby detect the presence of cancerous cells
XX
SQ Sequence 730 AA;
Query Match 99.9%; Score 3873; DB 4; Length 730;
Best Local Similarity 99.9%; Pred. No. 2.9e-293;
Matches 729; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MLFNSVLROPOLGVLNRGWSQYPLQSLLTGYQCSGNDEHTSYGETGVVPVPGCTFSSA 60
Db 1 MLFNSALRQPOLGVLNRGWSQYPLQSLLTGYQCSGNDEHTSYGETGVVPVPGCTFSSA 60
Qy 61 PMEHVLAVANEEGVRLYNTESQSFRRKCFKEMAHMNAVFDLAWVPGLKLVTAAGDQ 120
Db 61 PMEHVLAVANEEGVRLYNTESQSFRRKCFKEMAHMNAVFDLAWVPGLKLVTAAGDQ 120
Qy 121 TAKFWDVKAGELIGTCKGHQCSLKSVAFSKPKAVFCTGGRDGNIMVWTRCNKKGDFYR 180
Db 121 TAKFWDVKAGELIGTCKGHQCSLKSVAFSKPKAVFCTGGRDGNIMVWTRCNKKGDFYR 180
Qy 181 QVNOISGAHNTSDKOTPSKPKKQNSKGLAPSVDFQOSVTVVLFDENTLVSAGAVDGI 240
Db 181 QVNOISGAHNTSDKOTPSKPKKQNSKGLAPSVDFQOSVTVVLFDENTLVSAGAVDGI 240
Qy 241 KVDLRKNYTAAYROBPIASKSFYFGSSSTRKLGYSLLDSTGSTLFANCTDDNIYFMNM 300
Db 241 KVDLRKNYTAAYROBPIASKSFYFGSSSTRKLGYSLLDSTGSTLFANCTDDNIYFMNM 300
Qy 301 TGLKTSPIVAIFNGHONSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
Db 301 TGLKTSPIVAIFNGHONSTFYVKSLSLSPDDQFLVSGSSDEAAIYWKVSTPWQPTVLLGHS 360
Qy 361 QSVTVSVCCPSPDFTKIATCDDNTLKIWLNRGLKEKPGDKLSTVGWASQKKESRPG 420
Db 361 QSVTVSVCCPSPDFTKIATCDDNTLKIWLNRGLKEKPGDKLSTVGWASQKKESRPG 420
Qy 421 VVTSSQSTPAKAPRVKCNPSNPSAAACAPSCAGDLPLPSNTPTFSIKTSPAKARSP 480
Db 421 VVTSSQSTPAKAPRVKCNPSNPSAAACAPSCAGDLPLPSNTPTFSIKTSPAKARSP 480
Qy 481 NRRGSVSVSPKPPSFKMSIRNWNVTRTPSSSPPIPPASETKIMSPKALIPVSQKSSQ 540
Db 481 NRRGSVSVSPKPPSFKMSIRNWNVTRTPSSSPPIPPASETKIMSPKALIPVSQKSSQ 540
Qy 541 AEACSESRNRVKRRLLDSSCLSVKQKVCVSCNCVTELDGQVENLHLDLCLAGNQBDSLK 600
Db 541 AEACSESRNRVKRRLLDSSCLSVKQKVCVSCNCVTELDGQVENLHLDLCLAGNQBDSLK 600
Qy 601 DSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLPLRPGEGSEVMVGKENSSENK 660
Db 601 DSLGPTKSSKIEGAGTSISEPPSPISPYASESCGTLPLPLRPGEGSEVMVGKENSSENK 660
Qy 661 NWLLAAAKRAENPSPRSPSSQTPNSRRQSGTKLPSPTITPSSMRKICTYFHRKSQED 720
Db 661 NWLLAAAKRAENPSPRSPSSQTPNSRRQSGTKLPSPTITPSSMRKICTYFHRKSQED 720
Qy 721 FCGPEHSTEL 730
Db 721 FCGPEHSTEL 730

RESULT 9
ADQ97697
ID ADQ97697 standard; protein; 673 AA.
XX
AC ADQ97697;
XX
DT 07-OCT-2004 (first entry)
```


XX Human cancer associated sequence HP1-10-029, SEQ ID 674.
DE Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Human.
XX Homo sapiens.
XX WO2004060304-A2.
PN 22-JUL-2004.
XX 22-DEC-2003; 2003WO-US041389.
XX 27-DEC-2002; 2002US-00330773.
XX (SAGR-) SAGRES DISCOVERY INC.
PA Morris DW, Malandro MS;
PI WPI; 2004-543781/52.
XX New isolated cancer associated nucleic acids comprising at least 10
PT contiguous nucleotides, useful for diagnosing, preventing and/or treating
PT cancers such as leukemia and lymphoma.
XX Claim 1; SEQ ID NO 674; 199pp; English.
XX The present invention relates to cancer associated sequences (ADQ97025-
CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or
CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 673 AA;
SQ
Query Match 90.4%; Score 3503.5; DB 8; Length 673;
Best Local Similarity 91.8%; Pred. No. 28-264; 2; Indels 57; Gaps 2;
Matches 670; Conservative 1; Mismatches 2;
QY 1 MLFNSVLRQPLQGLVLRNGSSQVPLQSLLTGYQCSGNDEHTSYGTVGVPVPPFGCTFSSA 60
DB 1 MLFNSVLRQPLQGLVLRNGSSQVPLQSLLTGYQCSGNDEHTSYGTVGVPVPPFGCTFSSA 60
QY 61 PNMHVLAVANBEGFVRLYNTSSQSPRKKCFKEMAHNAVFDLAWVPGELKLVTAAGDQ 120
DB 61 PNMHVLAVANBEGFVRLYNTSSQSPRKKCFKEMAHNAVFDLAWVPGELKLVTAAGDQ 120
QY 121 TAKFWDVKAGELIGTCKGHQCSLKSVAFSKPEKAVFCTGGRDGNIMVWDTCKNKKDGEVR 180
DB 121 TAKFWDVKAGELIGTCKGHQCSLKSVAFSKPEKAVFCTGGRDGNIMVWDTCKNKKDGEVR 180
QY 181 QVNQISGAHNTSDKQTPSPKPKKQNSKGLAPSVDFQSQVTVVLFDQNTLVSAGAVDGI 240
DB 157 -----DFQSQVTVVLFDQNTLVSAGAVDGI 183
QY 241 KWDLRKNYAVRQEPFIASKGFYLPQSGSTRKLGYSLLIDSTGSTLFANCTDDNIYMFNM 300
DB 184 KWDLRKNYAVRQEPFIASKGFYLPQSGSTRKLGYSLLIDSTGSTLFANCTDDNIYMFNM 243
QY 301 TGLKTSFPAIFNGHONSTFVKSSLSPPDQFLVSGSSDEAAYTWKYSTPWPQPTVLLGHS 360
DB 244 TGLKTSFPAIFNGHONSTFVKSSLSPPDQFLVSGSSDEAAYTWKYSTPWPQPTVLLGHS 303
QY 361 QEVTSVCMCPSPDTFKIATCDNDTLKIWRNLGRLEKPGDKLSTVGMASQKKESRPG 420
DB 304 QEVTSVCMCPSPDTFKIATCDNDTLKIWRNLGRLEKPGDKLSTVGMASQKKESRPG 363
QY 421 VVTSSQSTPAKAPRVKCNPSNSSPSAACPACAGDLPLPSNTPFTSKTSPAKARSP 480
DB 364 VVTSSQSTPAKAPRVKCNPSNSSPSAACPACAGDLPLPSNTPFTSKTSPAKARSP 423
QY 481 NRRGSVSVSPKPPSSFKMSIRNWTPTPSSSPPIPPASETKIMSPKALIPVSKSSQ 540

DB 424 NRRGSVSVSPKPPSSFKMSIRNWTPTPSSSPPIPPASETKIMSPKALIPVSKSSQ 483
QY 541 AEACSESRNRVRRLLDSSCLSESVKQKCVKSCNCTVTELDGOVENLHLDLCCLAGNOEDLSK 600
DB 484 AEACSESRNRVRRLLDSSCLSESVKQKCVKSCNCTVTELDGOVENLHLDLCCLAGNOEDLSK 543
QY 601 DSLGPTKSSKIEGAGTISIEPPSPISPYASBSCGTLPLRPGCGGSEMVGKENSSENK 660
DB 544 DSLGPTKSSKIEGAGTISIEPPSPISPYASBSCGTLPLRPGCGGSEMVGKENSSENK 603
QY 661 NWLLAMAARKAENPSPSPSSQTPNSRRQSGKTLPSVTTTPSSMRKICITYFHRKSQED 720
DB 604 NWLLAMAARKAENPSPSPSSQTPNSRRQSGKTLPSVTTTPSSMRKICITYFHRKSQED 663
QY 721 FCGPEHSTEL 730
DB 664 FCGPEHSTEL 673
RESULT 10
ABU09610
ID ABU09610 standard; protein; 729 AA.
XX AC ABU09610;
XX DT 01-JUL-2003 (first entry)
XX Mouse retinoic acid-regulated nuclear matrix-associated protein.
DE Nucleic acid identification; T cell receptor; T cell; tumour antigen;
KW antigen presenting cell; HLA; nucleic acid library; ELISPOT assay;
KW enzyme-linked immunospot; T cell receptor binding epitope;
KW cancer associated antigen; ramp; mouse; polytope;
KW retinoic acid-regulated nuclear matrix-associated protein.
XX OS Mus musculus.
XX US2003003485-A1.
XX 02-JAN-2003.
XX 14-MAY-2002; 2002US-00145396.
XX 15-MAY-2001; 2001US-0291125P.
XX (LUDW-) LUDWIG INST CANCER RES.
XX Uenaka A, Nakayama E;
XX WPI; 2003-416593/39.
XX Identifying nucleic acids encoding epitopes that bind T cell receptors on
PT T cells when presented by human leukocyte antigen by using an ELISPOT
PT assay that detects factors secreted by T cell in response to the receptor
PT binding.
XX Example 6; Fig 13; 51pp; English.
XX The invention describes a method of identifying nucleic acid encoding an
CC epitope that binds to a T cell receptor on T cell. The method involves
CC coculturing antigen presenting cells that contain a nucleic acid library
CC and express HLA molecules presenting epitopes, with a T cell having a T
CC cell receptor that binds the epitope so that the T cell receptor binds
CC the epitope encoded by the nucleic acid library, detecting a factor
CC secreted by the T cell in response to T cell receptor binding, using
CC ELISPOT (enzyme-linked immunospot) assay, and correlating factor
CC secretion with presence of a nucleic acid encoding the epitope. The
CC method is useful for identifying a nucleic acid molecule encoding the
CC epitope that specifically binds to a T cell receptor on a T cell when
CC presented by an HLA molecule. A second method described in the invention
CC is useful for identifying antigens that specifically bind to a T cell
CC receptor on a T cell. The method is useful for identifying cancer

CC associated antigens. This is the amino acid sequence of the mouse
CC retinoic acid-regulated nuclear matrix-associated protein (ramp) peptides
CC of which can be combined with peptides of one or more other cancer-
CC associated antigens to form polytopes
XX
SQ Sequence 729 AA;
Query Match 89.7%; Score 3479.5; DB 6; Length 729;
Best Local Similarity 89.2%; Pred. No. 1.7e-262;
Matches 651; Conservative 36; Mismatches 42; Indels 1; Gaps 1;
QY 1 MLFNSVLRQPLGVLNRGNSQYPLQSLLTGYQCSGNDHTSYGETGVPPVPPGCTFSSA 60
DB 1 MLFNSVLRQPLGVLNRGNSQYPLQSLLTGYQCSGNDHTSYGETGVPPVPPGCTFCTA 60
QY 61 PNMEHLAVANEGFVRLYNTESSQSRKCKFKEMAHNAVFDLAWVPGBELKLVTAAGDQ 120
DB 61 PSMEHLAVANEGFVRLYNTESSQSRKCKFKEMAHNAVFDLAWVPGBELKLVTAAGDQ 120
QY 121 TAKFMDVKAGELGTCKGHQCSLKSVAFSKPEKAVFCTGGRCNIMVWDRCKNKGDFYR 180
DB 121 TAKFMDVRAGELGTCKGHQCSLKSVAFPRQKAVFSTGGRCNIMVWDRCKNKGDFYR 180
QY 181 QVNVQISGAHTSDKQTPSKPKKQNSKGLAPSVDFQCSVTTVLFDQENTLVSAGVDGII 240
DB 181 QVNVQISGAHTADKQTPSKPKKQNSKGLAPVDSQSVTVLFDQENTLVSAGVDGII 240
QY 241 KVDLRKNYTAHQEPIASKSFYPSGSTRKGLYSSILDSTGSTLPANCTDDNIYFNM 300
DB 241 KVDLRKNYTAHQEPIASKSFYPSGSTRKGLYSSILDSTGSTLPANCTDDNIYFNT 300
QY 301 TGLKTSFVALFNHONSTFVVKSSLSLDDQDFLVSGSSDEAAYTWKYSTPWQPTVLLGHS 360
DB 301 TGLKTSFVAVFNHONSTFVVKSSLSLDDQDFLVSGSSDEAAYTWKVSMPHPPPTVLLGHS 360
QY 361 QEVTSVCMWCPSTFKIATCDDNTLKIWLNRGLGEEKPGDGLSTVGWASQKKESRPGI 420
DB 361 QEVTSVCMWCPSTFKIATCDDNTLKIWLNRGLGEEKPGDGLSTVGWASQKKESRPGI 420
QY 421 VVTSSQSTPAKAPRVKCNPSNPSAACAAPCAGDLPLPSNTPFTSIKTPAKARSP 480
DB 420 VVTSSQSTPAKAPRAKSPSPSPSPSAACTPSCAGDLPLPSTPTFSVKTTPATTRSSV 479
QY 481 NREGSVSSVPKPPSPKSIKRWVTRTPSSSPITPPASETKIMSPRKALIPVSOKSSQ 540
DB 480 SRRGSISSVPKPLSFKWLSRNVTRTPSSSPVTPPASETKISSPRKALIPVSOKSSQ 539
QY 541 AEACSESRNRVRRRLDSSCLSEVKQKVCNCVTELDGOVENLHLDLCLAGNQBDLSK 600
DB 540 ADACSESRNRVRRRLDSSCLSEVKQKVCNCVTELDGOAELRLDLCLSGTQEVLSQ 599
QY 601 DSLGPTKSSKIEGAGTSISPPSPPIGPYASESCGTPLPLRPGEGSEVMGKNSPENK 660
DB 600 DSEGTPTKSSKIEGAGTSISPPSPVSPYASEGCGPLPLRPGEGSEVMGKNSPENK 659
QY 661 NWLLAAKKAENPSP 720
DB 660 NWLLAAKKAENPSP 719
QY 721 FCGPEHSTEL 730
DB 720 FCGPEHSTEL 729
RESULT 11
ADQ97694
ID ADQ97694 standard; protein; 786 AA.
XX AC ADQ97694;
XX DT 07-OCT-2004 (first entry)
XX DE Mouse cancer associated sequence MP10-029, SEQ ID 671.

XX Cytostatic; Gene Therapy; cancer; leukemia; lymphoma; Mouse.
KW Mus musculus.
OS
PN WO2004060304-A2.
XX
PD 22-JUL-2004.
XX
PF 22-DEC-2003; 2003WO-US041389.
XX
PR 27-DEC-2002; 2002US-00330773.
XX
PA (SAGR-) SAGRES DISCOVERY INC.
XX
PI Morris DW, Malandro MS;
XX
DR WPI; 2004-543781/52.
XX
PT New isolated cancer associated nucleic acids comprising at least 10
PT contiguous nucleotides, useful for diagnosing, preventing and/or treating
PT cancers such as leukemia and lymphoma.
XX
PS Claim 1; SEQ ID NO 671; 199pp; English.
XX
CC The present invention relates to cancer associated sequences (ADQ97025-
CC ADQ98004). The sequences are useful for the diagnosis, prevention and/or
CC treatment of cancer, such as leukemia and lymphoma. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 786 AA;
Query Match 75.1%; Score 2913.5; DB 8; Length 786;
Best Local Similarity 70.5%; Pred. No. 3.2e-218;
Matches 570; Conservative 34; Mismatches 46; Indels 159; Gaps 5;
QY 1 MLFNSVLRQPLGVLNR-----GWSSQYPLQSLLTGYQCSGND 17
DB 39 MLFNSVLRQPLGVLNRGAPWIAADCGPGLRPRERIRASEDRRTSCLGSPRTLTFTVI 98
QY 18 -----GWSSQYPLQSLLTGYQCSGND 38
DB 99 ELFSERRPKLGRSVASRAPGCLRQYRALTRCTLRVCGNSHYPQLSLLSYQNCND 158
QY 39 EHTSYGETGVPPVPPGCTFSSAPNMEHLAVANEGFVRLYNTESSQSRKCKFKEMAHW 98
DB 159 EHTSYGETGVPPVPPGCTFCTAPSMEHLAVANEGFVRLYNTESSQSRKCKFK----- 212
QY 99 NAVFDLAWVPGBELKLVTAAGDQTAQFMDVKAGELGTCKGHQCSLKSVAFSKPEKAVFCT 158
DB 213 ----- 212
QY 159 GGRDGNIMVWDRCKNKGDFYQVQVNIQISGAHTSDKQTPSKPKKQNSKGLAPSVDFQOS 218
DB 213 GGRDGNIMVWDRCKNKGDFYQVQVNIQISGAHTADQKTPSKPKKQNSKGLAPAVDSQOS 272
QY 219 VTVLFDQENTLVSAGVDGIIKVDLRKNYTAHQEPIASKSFYPSGSTRKGLYSSLI 278
DB 273 VTVLFDQENTLVSAGVDGIIKVDLRKNYTAHQEPIASKSFYPSGSTRKGLYSSLI 332
QY 279 LDSTGSTLPANCTDDNIYFNMVTLGKTSFVALFNHONSTFVVKSSLSLDDQDFLVSGSSD 338
DB 333 LDSTGSTLPANCTDDNIYFNMVTLGKTSFVALFNHONSTFVVKSSLSLDDQDFLVSGSSD 392
QY 339 EAAVTKWSTPWQPTVLLGHSQEVTSVCMWCPSTFKIATCDDNTLKIWLNRGLGEEKP 398
DB 393 EAAVTKWSTPWQPTVLLGHSQEVTSVCMWCPSTFKIATCDDNTLKIWLNRGLGEEKP 452
QY 399 GGDKLSTVGWASQKKESRPGI VVTSSQSTPAKAPRVKCNPSNPSAACAAPCAGDL 458
DB 453 -GDKHSIVGWTSQKKESRPGI VVTSSQSTPAKAPRAKSPSPSPSPSPSPSPSPSPSPSP 511

```
QY 459 PLPSNTPPTSIIKTSIPAKARSPINRRGSSVSSGPKPPSSPKMSIRNWNVTTPSSSPPTTPP 518
Db 512 PLPSSTPTFESVTKTTPATTRSSVSRGSISSGPKPLSSPKMSLRNWNVTTPSSSPPTTPP 571
QY 519 ASETKIMSPRKALIPVSQKSSQAEACSESRNRVKRRLDSSCLESVKQKCVKSCNCVTELD 578
Db 572 ASETKISSPRKALIPVSQKSSQADACSESRNRVKRRLDSSCLESVKQKCVKSCNCVTELD 631
QY 579 GQVENLHLDCLCLAGNQEDLSKSLGPTKSSKIEGAGTISBPSPSPISPYASESCGTLP 638
Db 632 GQAESLRDLCLCLSGTQEVLSQDSEGTSSKTEGAGTISBPSPSPVYASEGCGPLPL 691
QY 639 PLRPGCEGSEMGKENSSENKNWLLAMAAXKAENPSPRSPSSQTPNRRRSGKTLRSP 698
Db 692 PLRPGCEGSEMGKENSSENKNWLLAIAAKKAENSPRSPSSQTPNRRRSGKTLRSP 751
QY 699 VTITPSSMRKIC----TYFHRKSEQEDPCG 723
Db 752 STLA-----VGNHVQLVDAGIKCEFCG 774

RESULT 12
ID AAB92779 standard; protein; 434 AA.
XX
AC AAB92779;
XX
AC AC
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:11264.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000BP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 8; SEQ ID NO 11264; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesizing 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
```

```
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
XX
SQ Sequence 434 AA;
Query Match 59.1%; Score 2293; DB 4; Length 434;
Best Local Similarity 100.0%; Pred. No. 4.2e-170;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 297 MFNMTGLKTSVPAIFNGHONSTFYVKSSLSLSDPDQFLVSGSSDEAAAYIWKVSTPWPQPTVL 356
Db 1 MFNMTGLKTSVPAIFNGHONSTFYVKSSLSLSDPDQFLVSGSSDEAAAYIWKVSTPWPQPTVL 60
QY 357 LGHSQEVTSVCWCPDFTKIATCSDDNLTAKIWRNLNRGLEEKPGDKLSTVGWASQKKES 416
Db 61 LGHSQEVTSVCWCPDFTKIATCSDDNLTAKIWRNLNRGLEEKPGDKLSTVGWASQKKES 120
QY 417 RPLGLVTTSSOSTPAKAPRVKCNPSNSPSSAAACAPSCAGDLPLPSNTPPTSIIKTSIPAKA 476
Db 121 RPLGLVTTSSOSTPAKAPRVKCNPSNSPSSAAACAPSCAGDLPLPSNTPPTSIIKTSIPAKA 180
QY 477 RSPINRRGSSVSPKPPSSPKMSIRNWNVTTPSSSPPTTPPASSTKIMSPRKALIPVSO 536
Db 181 RSPINRRGSSVSPKPPSSPKMSIRNWNVTTPSSSPPTTPPASSTKIMSPRKALIPVSO 240
QY 537 KSSQAEACSESRNRVKRRLDSSCLESVKQKCVKSCNCVTELDGQVENLHLDCLCLAGNQ 596
Db 241 KSSQAEACSESRNRVKRRLDSSCLESVKQKCVKSCNCVTELDGQVENLHLDCLCLAGNQ 300
QY 597 DLSKDSLGPFTKSSKIEGAGTISBPSPSPISPYASESCGTLPPLRPGCEGSEMGKENS 656
Db 301 DLSKDSLGPFTKSSKIEGAGTISBPSPSPISPYASESCGTLPPLRPGCEGSEMGKENS 360
QY 657 PENKNWLLAMAAXKAENPSPRSPSSQTPNRRRSGKTLPSVTTTPSSMRKICTYFHRK 716
Db 361 PENKNWLLAMAAXKAENPSPRSPSSQTPNRRRSGKTLPSVTTTPSSMRKICTYFHRK 420
QY 717 SQEDFCGPEHSTEL 730
Db 421 SQEDFCGPEHSTEL 434

RESULT 13
ADM20092
ID ADM20092 standard; protein; 332 AA.
XX
AC ADM20092;
XX
XX
DT 20-MAY-2004 (first entry)
XX
DE Protein encoded by novel human channel/transporter gene #164 clone 2.
XX
KW immunosuppressive; antiarthritic; antirheumatic; antiproliferative;
KW cytotstatic; cardant; vasotropic; cerebroprotective; nootropic;
KW neuroprotective; antibacterial; virucide; fungicide; opthalmological;
KW gene therapy; channel/transporter protein; rheumatoid arthritis;
KW neoplasm; cardiac arrest; cerebrovascular disorder; cerebral ischemia;
KW angiogenesis; nervous system disorder; Alzheimer's disease;
KW ocular disorder; corneal infection; wound healing;
KW epithelial cell proliferation; skin aging; sunburn; transplantation;
KW chemotaxis; food additive.
XX
OS Homo sapiens.
```

PN WO200154472-A2.
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US0001307.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205115P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214986P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225113P.
PR 14-AUG-2000; 2000US-0225114P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231142P.
PR 08-SEP-2000; 2000US-0231143P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 23-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA Rosen CA, Barash SC, Ruben SM;
XX
PI

PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234224P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234978P.
PR 26-SEP-2000; 2000US-0234984P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-02556719P.

PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-476159/51.
XX N-PSDB; ADM19367.
PT Isolated nucleic acid molecule encoding a channel/transporter protein is
PT used in preventing, treating or ameliorating a medical condition.
XX
XX Claim 11; SEQ ID NO 653; 809pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule encoding a
CC channel/transporter protein or sequences at least 95% identical to a
CC these. The nucleic acids and proteins encoded by them are used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
CC in diagnosing a pathological condition or susceptibility to a
CC pathological condition. The antibodies to the proteins can also be used
CC in alleviating symptoms associated with the disorders and in diagnostic
CC immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
CC (ELISA). Disorders which are diagnosed or treated include autoimmune
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
CC neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
CC arrest, cerebrovascular disorders e.g. cerebral ischemia, angiodysplasia,
CC nervous system disorders e.g. Alzheimer's disease, infections caused by
CC bacteria, viruses and fungi and ocular disorders e.g. corneal infection.
CC The polypeptides can also be used to aid wound healing and epithelial
CC cell proliferation, to prevent skin aging due to sunburn, to maintain
CC organs before transplantation, for supporting cell culture of primary
CC tissues, to regenerate tissues and in chemotaxis. The polypeptides can
CC also be used as a food additive or preservative to increase or decrease
CC storage capabilities. This sequence corresponds to a protein of the
CC invention.
XX
SQ Sequence 287 AA;
Query Match 36.4%; Score 1413; DB 4; Length 287;
Best Local Similarity 100.0%; Pred. No. 1.3e-101;
Matches 266; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 109 GELKLVTAAGDQTAKEFDVKAGELIGTCCKGHCSSLKSVAFSEKAVFTGGRDGNIMYW 168
Db 3 GELKLVTAAGDQTAKEFDVKAGELIGTCCKGHCSSLKSVAFSEKAVFTGGRDGNIMYW 62
Qy 169 DTRCNKKGDFYRQVQVQISGAHNTSDKQTPSKPKKQNSKGLAPSDVDFQOSVTVLFPDEN 228
Db 63 DTRCNKKGDFYRQVQVQISGAHNTSDKQTPSKPKKQNSKGLAPSDVDFQOSVTVLFPDEN 122
Qy 229 TLVSAGAVDGI1KKWDLRKNYTAIRQEP1ASKSFYLPYSGSTRKLGYSIILDSGTGTLFA 288
Db 123 TLVSAGAVDGI1KKWDLRKNYTAIRQEP1ASKSFYLPYSGSTRKLGYSIILDSGTGTLFA 182
Qy 289 NCTDDNIYMFNMTGLKTSVAIFNGHONSTFVYKSLSPDDQFLVSGSDEAAYIKVST 348
Db 183 NCTDDNIYMFNMTGLKTSVAIFNGHONSTFVYKSLSPDDQFLVSGSDEAAYIKVST 242
Qy 349 PWQPPTVLLGHQSEVTSVCWCFSDFT 374
Db 243 PWQPPTVLLGHQSEVTSVCWCFSDFT 268
RESULT 15
AAB92715

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2006, 05:53:24 ; Search time 13047 Seconds
(without alignments)
12133.642 Million cell updates/sec

Title: US-10-726-160-1
Perfect score: 2831
Sequence: 1 ggcacgagcgagggttgag.....atgaaaaaaaaaaaaaaaaa 2831

Scoring table: IDENTITY NUC
Gapop 10,0 , Gapext 1.0

Searched: 48236798 seqs, 27959665780 residues

Total number of hits satisfying chosen parameters: 96473596

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:
2: gb_est3:
3: gb_est4:
4: gb_est5:
5: gb_est6:
6: gb_hc:
7: gb_est2:
8: gb_est7:
9: gb_est8:
10: gb_est9:
11: gb_gsa1:
12: gb_gsa2:
13: gb_gsa3:
14: gb_gsa4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1814	64.1	4205	6	AK166351	Mus muscu
2	1813.2	64.0	4180	6	AK133177	Mus muscu
3	1802.2	63.7	2350	6	AK054412	Mus muscu
4	1800.6	63.6	4179	6	AK164249	Mus muscu
5	1777.8	62.8	4201	6	AK169601	Mus muscu
6	1758	62.1	4228	6	AK154594	Mus muscu
7	1627.4	57.5	1923	14	DQ043568	Homo sapi
8	1539.2	54.4	3655	6	AK161401	Mus muscu
9	1451.8	51.3	1923	14	DQ043569	Pan trogl
10	996.2	35.2	1106	1	AL557830	AL557830
11	992.8	34.5	998	1	AL553783	AL553783
12	976.4	34.5	1042	1	AL576790	AL576790
13	927.2	32.8	1075	4	BX416133	BX416133
14	890	31.4	4308	6	AK164999	Mus muscu
15	885.8	31.3	949	3	BQ431493	AGENCOURT
16	881.6	31.1	1014	3	BM800170	AGENCOURT
17	874.6	30.9	907	4	BX462391	BX462391
18	856.4	30.3	905	1	AU123749	AU123749
19	852.2	30.1	958	1	AL558302	AL558302

20	851.4	30.1	1128	2	BM458369	AGENCOURT
21	847.4	29.9	914	3	BUI76525	AGENCOURT
22	845.8	29.9	962	1	AL580551	AL580551
23	845	29.8	929	3	BUI508613	AGENCOURT
24	824.2	29.1	899	4	CA488488	AGENCOURT
25	820.4	29.0	922	1	AL558817	AL558817
26	816.8	28.9	924	10	DV918737	DV918737
27	804.4	28.4	853	1	AU124656	AU124656
28	799.4	28.2	883	3	BQ233905	AGENCOURT
29	794	28.0	902	3	BQ34727	AGENCOURT
30	788.2	27.8	866	4	EX435948	EX435948
31	786	27.8	804	8	CA164038	AGENCOURT
32	760.6	26.9	820	5	CD651902	CD651902
33	755	26.7	936	2	BI257561	BI257561
34	751.6	26.5	1012	3	BM557720	AGENCOURT
35	749	26.5	843	9	DA736257	DA736257
36	733.4	25.9	1107	2	BM468859	AGENCOURT
37	731	25.8	731	9	CA786033	CA786033
38	729.4	25.8	902	3	BU521670	AGENCOURT
39	725	25.6	727	8	CN362845	CN362845
40	719.2	25.4	847	8	CV556506	CV556506
41	712.2	25.2	962	7	BE746556	BE746556
42	708	25.0	708	8	CN362847	CN362847
43	707.4	25.0	763	2	BG171017	BG171017
44	701.4	24.8	981	7	BE742123	BE742123
45	700.8	24.8	711	9	DN996329	DN996329

ALIGNMENTS

RESULT 1
AK166351
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS

AK166351 4205 bp mRNA linear HTC 21-SEP-2005
Mus musculus mammary gland RCB-0526 Jyg-MC(A) cDNA, RIKEN
full-length enriched library, clone:G830048A21 product:L2D7L
protein, full insert sequence.
AK166351.1 GI:74147704
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
10349636
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
11042159
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
11076861
4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,

Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bona, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.

RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium

Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)

11217851

5

CONSRMT
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Chochia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reid, J. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

FANTOM Consortium

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420 (6915), 563-573 (2002)

12466851

6

CONSRMT
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M. C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shimokawa, K., Bajic, V. B., Brenner, S. E., Batalov, S., Forrest, A. R., Zavolan, M., Davis, M. J., Wilming, L. G., Aidinis, V., Allen, J. E., Ambesi-Impombato, A., Anweiler, R., Aturaliya, R. N., Bailey, T. L., Bansal, M., Baxter, L., Belsel, K. W., Bersano, T., Bono, H., Chalk, A. M., Chiu, K. P., Choudhary, V., Christoffels, A., Clutterbuck, D. R., Crowe, M. L., Dalla, E., Dalrymple, B. P., de Bono, B., Della Gatta, G., di Bernardo, D., Down, T., Engstrom, P., Fagioli, M., Faulkner, G., Fletcher, C. F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgii-Hemming, P., Gingeras, T. R., Gojobori, T., Green, R. E., Gustincich, S., Harbers, M., Hayashi, Y., Hensch, T. K., Hirokawa, N., Hill, D., Hummel, L., Iacono, M., Ikeo, K., Iwama, A., Ishikawa, T., Jakt, M., Kanapin, A., Katoh, M., Kawasawa, Y., Kelsso, J., Kitamura, H., Kitano, H., Kollias, G., Krishnan, S. P., Kruger, A., Kummerfeld, S. K., Kurochkin, I. V., Lareau, L. F., Lazarevic, D., Lipovich, L., Liu, J., Ljun, S., McWilliam, S., Madan, Babu, M., Madera, M., Marchionni, L., Matsuda, H., Matsuzawa, S., Miki, H., Mignone, F., Miyake, S.,

Morris, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakauchi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Parg, K. C., Pavan, W. J., Pavese, G., Pesole, G., Petrovsky, N., Piazza, S., Reed, J., Reid, J. F., Ring, B. Z., Ringwald, M., Rost, B., Ruan, Y., Salzberg, S. L., Sandelin, A., Schneider, C., Schonbach, C., Sekiguchi, K., Semple, C. A., Seno, S., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sinclair, B., Sperling, S., Stupka, E., Sugiyara, K., Sultana, R., Takanaka, Y., Taki, K., Tammoja, K., Tan, S. L., Tang, S., Taylor, M. S., Tegner, J., Teichmann, S. A., Ueda, H. R., van Nimwegen, E., Verardo, R., Wei, C. L., Yagi, K., Yamanishi, H., Zabarovsky, E., Zhu, S., Zimmer, A., Hide, W., Bult, C., Grimmond, S. M., Teasdale, R. D., Liu, E. T., Brusic, V., Quackenbush, J., Wahlestedt, C., Mattick, J. S., Hume, D. A., Kai, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori-Katayama, M., Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kato, T., Kawaji, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondo, S., Konno, H., Nakano, K., Ninomiya, N., Nishio, T., Okada, M., Plessey, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Watahiki, A., Okamura-Oho, Y., Suzuki, H., Kawai, J. and Hayashizaki, Y.

FANTOM Consortium

The transcriptional landscape of the mammalian genome
Science 309 (5740), 1559-1563 (2005)

16141072

7

CONSRMT
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS

Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yagi, C. C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T., Pang, K. C., Hallinan, J., Mattick, J., Hume, D. A., Lipovich, L., Batalov, S., Engstrom, P. G., Mizuno, Y., Faghihi, M. A., Sandelin, A., Chalk, A. M., Mottagui-Tabar, S., Liang, Z., Lenhard, B. and Wahlestedt, C.

RIKEN Genome Exploration Research Group

Antisense transcription in the mammalian transcriptome
Science 309 (5740), 1564-1566 (2005)

16141073

8 (bases 1 to 4205)

CONSRMT
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS

Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hori, F., Iida, J., Imamura, K., Imotani, K., Itoh, M., Kanagawa, S., Kawai, J., Kojima, M., Konno, H., Murata, M., Nakamura, M., Ninomiya, N., Nishiyori, H., Nomura, K., Ohno, M., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shiraki, T., Tagami, M., Tagami, Y., Waki, K., Watahiki, A., Muratsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (14-APR-2004) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: genome-res@gs.riken.jp).
URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.
URL: <http://genome.gsc.riken.jp/>
URL: <http://fantom.gsc.riken.jp/>
Location/Qualifiers

1. 4205
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="FANTOM_DB:G830048A21"
/db_xref="taxon:10090"
/clone="G830048A21"
/cell_line="RCB-0526 JYG-MC(A)"
/tissue_type="mammary gland"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
123. 2312
/note="unnamed protein product; L2DTL protein (UniProt|Q8BW38, evidence: FASTA, 99.7%ID, 100%length, match=2187) putative"

CDS

FEATURES
source

		/codon_start=1 /protein_id="BAE38725.1" /db_xref="GI:74147705" /translation="MLFNSVLRLQPOLGVLRNWSHYPLQSLLSGVQCNDEHTSYG ETGVPPFGCTFCTAPSMHEHLIVANEELFVRLYNTEQTSKTKCFKEMAHNAVF DLAWPEGLKLVTAAGDQTAKEFMDVRAELMGTCGHQCSLKSVAFSKFQAVFTGG RDGNIIMIDTRCNKKDGFYRQVQNLISGAHNTADKTPSPKKQNSKGLAPAVDSQOS VTWLPDQNTLIVSAGAVGIIKMDLRKNYTAIROEPIASKFLYPGTSTRKLGYS LVLDSTGLTFANCTDNDIYFMNTGLKTSPIAVENGHQNSTFYVYKSSLSPDDQLIS GSDENAYIKWISMPHPPTVLLHGSQEVTSVCMCFSPDFIKIATCSDNLTILWRNL GLEEKPDHSLVIGTWSOKKEVKACPTVPSSOSTPAKAPRAKSPSSISSPSAACT PSCAGDPLPLSPSTPTSVKTTTTPATTRSSVSRGSISSVSPKPLSPFMSLRNWRTP SSSPPTVPASPTKISSPKALIPVQSKSQADACSESRNRVKRLRLDCLSLSVQKC VKSNCNVTELDQBSRLRLDCLSLQAQEVLSQDSSEPTKSSKTEGAGTSSIEBPPSPV		Query Match 64.1%; Score 1814; DB 6; Length 4205; Best Local Similarity 86.9%; Pred. No. 0; Matches 2018; Conservative 0; Mismatches 300; Indels 4; Gaps 2;	
QY	8	GC	GGGAGTTGGAGCGGATAACGATTGTTGTGTGAGAGCGCAACGTCGGATTTCCTGCTG 67		
DB	6	GC	GGGAGTTGGAGCGCAACCGCGCGGTGTGTGGAGGTTGCAGAGCGCGCTCTCTGTGG 65		
QY	68	AA	CTTGGAGGC-ATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATG 126		
DB	66	AA	CTTGGGCGGATCTTCTCGCGGTTCCCTCAGCTGAGCTTCTCTCAGGCCCAGATG 125		
QY	127	CT	CTTCAATTCGTGCTCGCCAGCCGACGCTTGGCGTCTGTGAGAAATGGAATGATGCTTCA 186		
DB	126	CT	CTTCAACTCGGTGCTCGCCAGCCGACGCTCGCGTCTGTGAGGAACGGGTGCTTCA 185		
QY	187	CA	ATACCTCTTCAATCCCTTCTGACTGTGTATCAGTCAGTCAGTGGTAAATGATGAACACAT 246		
DB	186	CA	TATACCTCTCAATCCCTTCTAAGTGGTATCAGTCGCAACTGTAAACGATGAACACAG 245		
QY	247	TC	TTATGGAGAAAACGAGGTCCTCCTCTTTTGGATGTACCTTCTCTTCTGCTCCC 306		
DB	246	TC	TTATGGAGAAAACGAGGTCCTCCTCTTTTGGATGTACCTTCTGTACCGCTCCC 305		
QY	307	AA	TATGGAACTACTAGCAGTTGCCAATGAAGAAGCTTTGTTTCAATTTGTTATATACACA 366		
DB	306	AG	TATGGAGCATATATTAGCAGTTGCTAATGAAGAAGCTTTTGTCAAGATATATATATACA 365		
QY	367	GA	TACAAAGTTTCAGAAAGAGTCTTCAAGAAATGATGGCTCACTCGGAATGCCGTC 426		
DB	366	GAG	TACAAATAGCAAAAAGACATGCTTCAAGAGTGGATGGCTCACTGGAAATGCTGTC 425		
QY	427	TT	TGACCTGGCCTGGGTTCTCGGTGAACCTTAACTTTGTTACAGCAGCAGGTGATCAAAACA 486		
DB	426	TT	TGACTTGGCCTGGGTCCTCGTGAACCTTAACTTTGTTACAGCAGCAGGTGATCAAAACA 485		
QY	487	GC	CAAAATTTTGGAGCTAAAGCTGTGTAGCTGATTGGAACTATGCAAAAGTCAATATGC 546		
DB	486	GC	CAAAATTTTGGAGCTTAAGAGCTGTGTAGCTTAAATGGGAACTATGCAAAAGGCCACAGTGC 545		
QY	547	AG	CTCAAGTCAGTTCCTTTTCTAAGTTTGAGAAAGCTGTTATTTCTGTACGGGTGGAAGA 606		
DB	546	AG	CTCAAGTCTGTAGCTTTTCCAAAGTTTCAAAAGCTGTGTTCTCTACAGGGGGGAGA 605		
QY	607	GA	TGGCAACATTTATGTCTGGGATACCAAGGTGCAACAAAAAGATGGGTTTTATAGGCAA 666		
DB	606	GAC	GCAACATTTATGATCTGGGACACCAAGGTGTAACAAAAAGATGGATTTTATAGACAA 665		
QY	667	GT	GAATCAATCAGTGGAGCTCACAATACCTCAGACAAGCAACCCCTTCAAAACCCCAAG 726		
DB	666	GT	GAATCAATCAGTGGAGCTCACAATATCTGAGACAAGCAACCCCTTCAAGGCCCAAG 725		
QY	727	AG	AAACACAGAAATCAAAGGACTTGTCTCTCTGTGGATTTCAGCAAAAGTGTACTGTG 786		
DB	726	AG	AAACAAAAATCAAAGGACTTGTCTCTCTGTGGATTCCCAGACAGATGTTTACTGTG 785		
QY	787	GT	CCTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATTAATCAAA 846		

786	GT	CCTCTTT	CAGGATGAGAAATACATTAGTCTCAGCAGGAGCGGTGGATGAATAATCAAA	845
847	GT	ATCGGATTT	ACGTAAGAAATTTATATCTCTTATCACAAGAACCCATAGCATCCCAAGTCT	906
846	GT	ATCGGATTT	TCGCGCAAGAAATTTACACTGCTTATCACAAGAACCCATAGCATCCCAAGTCT	905
907	TT	CCTGTAT	CCCCAGGTAGCAGCAGCTCGAAACCTTGGATATTCAAGTCTGATTTTGGATTCC	966
906	TT	CCTGTAT	CCCCAGGTACCAAGCTAGGATACTCGAGTTGGTTTGGTTTATAGACTCT	965
967	ACT	GCTCTA	CTTTATTTGCTAAATTTGCACAGACGATAACATCTACATGTTTAAATATGACT	1026
966	ACT	GCTCTA	CTTTATTTGCTAACTTCACAGATGACACATCTATATGTTTAAATATGACT	1025
1027	GG	TTTGA	GACTTCTCCAGTGGCTATTTTCAATGACACACAGAACTCTACCTTTTATGTA	1086
1026	GG	CTTAA	AGACTTCTCCGGTGGCTGTCTTCAATGACACACAGAACTCTACCTTTTATGTA	1085
1087	AA	ATCCAG	CCTTAGTCCAGATGACCAAGTTTTTGTAGTCAAGTGGCTCAAGTGATGAAGCTGCC	1146
1086	AA	ATCAAG	CTTTAGTCCAGATGACCAAGTTTTTAAATCAGTGGTTCAAGTGATGAAGCTGCC	1145
1147	TA	CATATG	GAAGGCTCTCACACCTGGCAACCTCTACTGTGCTCTCTGGGTCAATTTCTCAA	1206
1146	TA	CATTTG	GAAGGTTTCCATGCGCATCGCATCTCTACTGTGCTCTCTGGGTCAATTTCTCAA	1205
1207	GAG	GTCA	GCTGTGTGTGCTGCTCATCTGACTTTCAAAAGATTGCTACCTGTTCTGAT	1266
1206	GAG	GTCA	GCTGTGTGTGCTGCTCATCTGACTTTCAAAAGATTGCTACCTGTTCTGAT	1265
1267	GA	CAATAC	ACTAAAAATCTGGCGCTTGAATAGAGCTTTAGAGGAGAAAAACAGAGGTGAT	1326
1266	GAT	ATAC	ACTGAAATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACC---AGGTGAT	1322
1327	AA	ACTTTT	CCACGGTGGGCTCTCAGAAAGAAAAAGAGTCAAGACTCGGCTTAGTA	1386
1323	AA	ACATTT	CCATAGTGGGTTGGACCTCTCAGAAAGAAAAAGAGTCAAGACTCGGCTTAGTA	1382
1387	AC	AGTAA	CGAGTAGCAGAGTACTCTGCCAAGCCCCAGGGTAAAGTCAATCCATCC	1446
1383	AC	GGTAC	CAAGTAGCAGAGTACTCTGCCAAGCTCCCAGAGCCCAAGAGCAGTCCATCC	1442
1447	AA	TTCTT	CCCCCTCATCGCAGCTTGTGCCCCCAAGCTGTCTGTGGAGACCTCCCTCTTCT	1506
1443	AT	CTCT	CTCTCTCTCGTCTCAGAGCTTGTACTCGAGCTGTGCAGAGAGACCTCCCTCTTCT	1502
1507	TC	AAAT	ATCTCTTAATAAAACCTCTCTGCCAAGCCCGGTCTCCCATCAAC	1566
1503	TC	AAAT	ATCTCTTAATAAAACCTCTCTGCCAAGCCCACTCTCCGCAAGCCCGGTCTCT	1562
1567	AG	AAGAG	GTCTGTCTCTCGTCTCTCCAGCCACCTTTCATCTTCAAGATGTCGATT	1626
1563	AG	AAGAG	GTCTGTCTCTCGTCTCTCCAGCCACCTTTCATCTTCAAGATGTCGATT	1622
1627	AG	AAAT	CTGGTGACCCGAAACACTTCTCTCATCACCAACCCATCACTCCACTGCTTGGAG	1686
1623	AG	AAAT	CTGGTGACCCGAAACACTTCTCTCATCACCAACCCATCACTCCACTGCTTGGAG	1682
1687	AC	CAAG	ATCATGTCTCCAGAAAAAGCCCTTATTCCTGTGAGCCAGAGTCAATCCCAAGCA	1746
1683	AC	CAAG	ATCATGTCTCCAGAAAAAGCCCTTATTCCTGTGAGCCAGAGTCAATCCCAAGCA	1742
1747	GAG	GCTT	GCTCTGAGCTTAGAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGCTGGAG	1806
1743	GAT	GCTT	GCTCTGAGCTTAGAATAGAGTAAAGAGGAGGCTTGTGACTCAAGCTGCTGGAG	1802
1807	AG	TGTG	AAACAAAGTGTGTGAAGATTGTAACTGTGTGACTGTGACTTGTATGGCCAAAGTT	1866
1803	AG	TGTG	AAACAAAGTGTGTGAAGATTGTAACTGTGTGACTGTGACTTGTATGGCCAAAGTT	1862
1867	GAAA	ATCT	TTCATTGGATCTGTGCTCTGTGTGTAACACAGGAGAGCTTATGTATAGGAC	1926
1863	GAG	ATCT	TTCATTGGATCTGTGCTCTGTGTGTAACACAGGAGAGCTTATGTATAGGAC	1922

QY	1927	TCCTAGTCCCTACCAAAATCAAGCAAAATTAAGGAGCTGGTACCAGTATCTCAGAGCCT	1986
DB	1923	TCGAGGGTCTTACCAAAATCAAGCAAGACTGAAGGTGGCAAGCATCTCAGAACCT	1982
QY	1987	CGTCTCTCATCAGTCGCTATCCTTCAGAAAGCTGTGAAGCCTACCTCTTCTCTTTGAGA	2046
DB	1983	CCTTCTCTCTGTCAGTCTTATCTGCTTGAAGGCTGTGGACCACTGCTCTTCTCTTTGAGA	2042
QY	2047	CCTTGTGGAGAAAGGGTCTGAAATGGTAGGCAAGAGAAATAGTTCCTCCAGAGAAATAAAAAC	2106
DB	2043	CCTTGTGGAGAGGATCTGAGATGGTGGCGGAGAGNATAGTCTCTCCAGAGAAATGAAGAC	2102
QY	2107	TGGTTGTGGCCATGCGCAGCCAAAGCGGAAGGCTGAGATCCATCTCTCCAGGAAGTCCGCTCA	2166
DB	2103	TGGTTGTGGCCATGAGCAGCCAAAGCGGAGGAGAAAATTCATCCCGAAGAGTCCATCA	2162
QY	2167	TCCCAGACACCAATTCAGGAGAGAGAGCGGAGAAAGACATTCGCCAGCCGTCACCATC	2226
DB	2163	TCTCAGACACCCAGTTCAGAGGAGCAAAAGTGGAGAGAGCTCACCAGGCCGCTCACCATT	2222
QY	2227	AGCCCGAGCTCCATGAGGAAATCTGCACATACTTCCATAGAAAGTCCCGAGGAGCTTC	2286
DB	2223	ACTCCAGCTCCATGAGGAAGATATGTACATACTTTCGTAGAAGACTCAAGATGACTTC	2282
QY	2287	TGTGCTCTGAACACTCAACAGAAATATAGATCTTAATCTGA	2328
DB	2283	TGCAGTCTTGAACACTCAACTGAATATATAGATGCTAATCTGA	2324
RESULT 2			
AK133177			
LOCUS	AK133177	4180 bp mRNA linear	HTC 21-SEP-2005
DEFINITION	Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4931420L14 product:l2D7L protein, full insert sequence.		
ACCESSION	AK133177		
VERSION	AK133177.1	GI:74221710	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus	(house mouse)	
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1		
AUTHORS	Carninci, P., Shibata, Y., Hayata, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res.	10 (10), 1617-1630	(2000)
PUBMED	11042159		
AUTHORS	2		
	Carninci, P., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kusunagi, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, Y., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multipicapillary sequencer		
JOURNAL	Genome Res.	10 (11), 1757-1771	(2000)
PUBMED	11076861		
AUTHORS	4		
	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,		

Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Schenkl, P., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barab, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombauts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seva, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohetsuki, S. and Hayashizaki, Y.	
RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium	
FUNCTIONAL annotation of a full-length mouse cDNA collection	
Nature 409 (5821), 585-590 (2001)	
11217851	
5	
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusci, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maitais, L., Marchionni, L., McKenzie, L., Mik, H., Nagashima, T., Numa, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Sempie, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.	
FANTOM Consortium	
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
Nature 420 (6915), 563-573 (2002)	
12466851	
6	
Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shimokawa, K., Bajic, V.B., Brenner, S.E., Batalov, S., Forrest, A.R., Zavolan, M., Davis, M.J., Wilming, L.G., Aidinis, V., Allen, J.E., Ambesi-Impombato, A., Apweiler, R., Aturaliya, R.N., Bailey, T.L., Bansal, M., Baxter, L., Beisel, K.W., Bersano, T., Bono, H., Chalk, A.M., Chiu, K.P., Choudhary, V., Christoffels, A., Clutterbuck, D.R., Crowe, M.L., Dalla, E., Dalymple, B.P., de Bono, B., Della Gatta, G., di Bernardo, D., Down, T., Engstrom, P., Fagioli, M., Faulkner, G., Fletcher, C.F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgii-Hemming, P., Gingeras, T.R., Gojobori, T., Green, R.E., Gustincich, S., Harbers, M., Hayashi, Y., Hensch, T.K., Hirokawa, N., Hill, D., Huminecki, L., Iacono, M., Ikeo, K., Iwama, A., Ishikawa, T., Jakt, M., Kanapin, A., Katoh, M., Kawasawa, Y., Kelso, J., Kitamura, H., Kitano, H., Kollias, G., Krishnan, S.P., Kruger, A., Kummerfeld, S.K., Kurochkin, I.V., Lareau, L.F., Lazarevic, D., Lipovich, L., Liu, J., Liuni, S., McWilliam, S., Madan, Babu, M., Madera, M., Marchionni, L., Matsuda, H., Matsuzawa, S., Miki, H., Mignone, F., Miyake, S., Morris, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakachi, H.,	

Qy	851	GGGATTTCAGTAAGAAATTATATCTCTTATCGACAAGAACCCATAGCATCCAAGTCTTTCC	910
Db	841	GGGATTTCGGCAAGAAATTACACTGCTTATCGACAAGAACCCATAGCATCCAAGTCTTTCC	900
Qy	911	TGTACCAGGTPAGCAGCACTCGAAAACTTGGATATTCAAGTCTGATTTTGGATTCACATG	970
Db	901	TGTACCAGGTPACCAGCACTCGAAAAGTAGGATACTCGAGTTTGGTTTTAGACTCTACTG	960
Qy	971	GCTCTACTTATTGTCTTAATTGACACAGCATACATCTACATGTTTAAATATGACTGGGT	1030
Db	961	GCTCTACTTATTGTCTTAATGTCACAGATGACAACTATATGTTCAATATGACTGGGT	1020
Qy	1031	TGAAGACTTCTCCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTATGTAAAAAT	1090
Db	1021	TAAAGACTTCTCCGGTGGCTCTTCAATGGACACAGAACTCTACCTTTTATGTAAAAAT	1080
Qy	1091	CCAGCCTTAGTCCAGATGACAGATTTTTAGTCAGTGGCTCAAGTGTGAAGTGCCTACA	1150
Db	1081	CAAGCTTAGTCCAGATGACCACTTTTAAATCAAGTGTCAAGTGTGAAGTGCCTACA	1140
Qy	1151	TATGGAGGTTCTCACACCTTGGCAACCTCTACTGTGCTCTGGGTCAATCTCAAGAGG	1210
Db	1141	TTTGGAGGTTTCCATGCCATGGCACTCCTCTACTGTGCTCTGGGTCAATCTCAAGAGG	1200
Qy	1211	TCAGTCTGTGCTGGTGTCCATCTGACTTTCACAAAGATTGCTTACCTGTCTCGATGACA	1270
Db	1201	TCAGTCTGTGCTGGTGTCCATCAGACTTTCACCAAGATTGCAACTGTCTCGATGATA	1260
Qy	1271	ATACACTAAAAATCTGGCGCTTGAATAGAGCTTTAGAGGAAACCAAGGAGGTGATAAAC	1330
Db	1261	ATACACTGAAAATCTGGCGCTTGAATAGAGCCTTAGAGGAAACCAAGGAGGTGATAAAC	1317
Qy	1331	TTTCCAGGTGGTGGGCTCTCAGAAAGAAAAGAGTCAAGCTTGGCCTTAGTAACAG	1390
Db	1318	ATTCCATAGTGGTGGACCTCTCAGAAAGAAAAGAGTCAAAAGCTGCGCCAGTAACGG	1377
Qy	1391	TAAAGAGTGGCAGAGTACTCTCCCAAGGCCCCAGGGTAAAGTGCAATCCATCCAATT	1450
Db	1378	TACCAAGTAGCCAGAGTACTCTCCCAAGAGTCCCAAGGCCAAGAGCAGTCCATCCATCT	1437
Qy	1451	CTTCCCGTCAATCCGCACTGTGTGCCCAAGCTGTGTGGAGACCTCCCTCTCTCTTCAA	1510
Db	1438	CTCTCTCTGTGTCAGAGTGTACTCCGAGCTGTGCAGAGAGCTCCCTCTCTCTTCAA	1497
Qy	1511	ATACTCTACGTTCTTATTAACCTCTCTCTGCAAGGCCCGGTCTCCCATCAACAGAA	1570
Db	1498	GTACCCCAATCTCTAGTCAAAACCACTCTCTGCAAGGCCCGGTCTCTCAGTCAGCAGAA	1557
Qy	1571	GAGCTCTGTCTCTCGTCTCTCCCAAGCCACTTCACTTTTCAAGATGTCGATTAGAA	1630
Db	1558	GAGGCTCCATCTTCTGTGTCTCCCAAGCCACTTCACTTTTCAAGATGTCGATTAGAA	1617
Qy	1631	ACTGGGTGACCCGAACCTTCTCATCAACCAACCATCACTCCACCTGTCTCGAGACCA	1690
Db	1618	ACTGGGTGACCCGAACCTTCTCATCAACCACTGTCACTCCACCTGTCTCTGAGACAA	1677
Qy	1691	AGATCATGTCTCCAGAAAAAGCCCTTATCTGTGAGCCAGAGTCAATCCCAAGCAGAGG	1750
Db	1678	AGATCTCATCTCCAAGAAAAGCTTATCTCTGTGTGAGCCAGAGTCAATCAAGCCAGATG	1737
Qy	1751	CTTGCTCTGAGTCTAGAAAATAGATTAAGAGGAGCTTAGACTCAAGCTGTCTGGAGATG	1810
Db	1738	CTTGCTCTGAAATCTAGAAAATAGATTAAGAGGAGCTTGTGACTCAAGCTGTCTGGAGATG	1797
Qy	1811	TGAAACAAAAGTGTGTGAAGATTGTAACTGTGTGACTGTAGCTTGTATGGGCCCAAGTTGAAA	1870
Db	1798	TGAAACAAAAGTGTGTGAAGATTGTAACTGTGTGACTGTAGCTTGTATGGGCCCAAGCGAGA	1857
Qy	1871	ATCTTCATTTGGATCTGTGTGCTTGTGTTGTTAAACCAAGGAAGACCTTAGTAAGGACTCTC	1930
Db	1858	GTCCTCGTTTGGATCTGTGTGCTTGTGTTTCCGACCCAGGAAGTCTTAGGCCAAGACTCCG	1917

Qy	1931	TAGGTCTTACCAAAATCAAGCAAAATTTGAAGGAGCTGGTACCAAGATATCTCAGAGCCTCCGT	1999
Db	1918	AGGGTCTTACCAAAATCAAGCAAGACTGAAGGTGCTGGCACAAGCATCTCAGAACTCTCCCT	1977
Qy	1991	CTCTATCAGTCCGTATGCTTTCAGAAAGCTGTGGAAGCTACTCTCTCTTTTTCAGACCTT	2050
Db	1978	CTCCTGTCACTCCGTATGCTTTCAGAAAGCTGTGGAAGCTACTCTCTCTTTTTCAGACCTT	2037
Qy	2051	GTGGAGAAGGGTCTGAAATGGTATAGGCAAGAGAGAAATAGTTCCTCCAGAGAAATATAAACTGGT	2110
Db	2038	GTGGAGAGAGATCTGAGATGGTGGCAAGAGAGAAATAGTCTCCAGAGAAATAGAACTGGT	2097
Qy	2111	TGTTGGCCATGGCAGGCCAAACCGAAGCTGTAGAAATCCATCTCCACGAAGTCCCTATCCC	2170
Db	2098	TGTTGGCCATAGCAGGCCAAACCGAAGCTGTAGAAATCCATCTCCACGAAGTCCCTATCCC	2157
Qy	2171	AGACACCAATTCAGGAGACAGAGCGGAAGAGCATTTGCCAAGCCCGGTCCACCATCACGC	2230
Db	2158	AGACACCAATTCAGGAGACAGAGCGGAAGAGCATTTGCCAAGCCCGGTCCACCATCACGC	2217
Qy	2231	CCAGCTCCATGAGGAAATCTCGCACATATCTTCATAGAAAGTCCCAAGGAGGACTTCTGTG	2290
Db	2218	CCAGCTCCATGAGGAGATATGTACATCTTTCGTAGAAAGATCAAGATGACTTCTGTG	2277
Qy	2291	GTCTGAACACTCAACAGAAATATAGATTTAAATCTGA	2328
Db	2278	GTCTGAACACTCAACTGAATATAGATGCTAAATCTGA	2315
RESULT 3			
AK054412			
LOCUS			
DEFINITION			
AK054412	2350 bp	mRNA	linear HTC 02-SEP-2005
DESCRIPTION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			
LOCUS			
DEFINITION			
AK054412			

REFERENCE AUTHORS	5 The FANTOM Consortium, the RIKEN Genome Exploration Research Group Phase I and II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL REFERENCE	Nature 420, 563-573 (2002)
AUTHORS	6 RIKEN Genome Exploration Research Group, Genome Science Group (Genome Network Core Team) and the FANTOM Consortium.
TITLE	Antisense Transcription in the Mammalian Transcriptome
JOURNAL REFERENCE	Science 309, 1564-1566 (2005)
AUTHORS	7 The FANTOM Consortium, Riken Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group).
TITLE	The Transcriptional Landscape of the Mammalian Genome
JOURNAL REFERENCE	Science 309, 1559-1563 (2005)
AUTHORS	8 (bases 1 to 2350) Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Maezuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akai, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL REFERENCE	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: genome-resgsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/ Location/Qualifiers 1. 2350 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:E330023F01" /db_xref="taxon:10090" /clone="E330023F01" /sex="female" /tissue types="ovary" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="2 days pregnant adult" 105..2294 /notes="unnamed protein product; L2DTL PROTEIN (RA-REGULATED NUCLEAR MATRIX-ASSOCIATED PROTEIN) homolog [Homo sapiens] (SPTR Q9NZJ0, evidence: FASTA, 89.5%ID, 100%length, match=2187) putative" /codon_start=1 /protein_id="BAC35769.1" /db_xref="GI:26344225"
CDS	translation="MLPNSVLROPLQGLRWGSSHYPLQSLISGVQCNDBHTSYG ETGVPPFGCTCTAPSMHILAVANEGFVRLYNTESETQTSKCTCKEWMHWNVAF DLAWVGEKLVTAGDQFAKFWDRAGELMGCTKGHQSLSKSVAFPKQKAVFTSGG RDGNIMWDRNCNKGDFRQVQNIQSGAHNTADKQTPSPKKQNSGLAPVDSQGS VTVLQFDENLVSAGADGIIKVDLRKNYTAQROEPIASKSFLYPTGSTRKLGYSS LVLDSTGLTFANCTDNTYMFNMTGLKTSFPAVFNHQNSTFYVKSLSLSPDDQFLIS GSSDEAAYIKWYMPHPPTVLLGHSEQVTSVCWCPSPDFTKIATCSDDDNTLKIWLNR
ORIGIN	GLEKPGDKHSIVGWTSSQKKKEVKACPVTVPSQSSQTPAKAPRAKSPSSISSAACT PSCAGPLPSPPTPTSVKTTTATTRSSVSSVSPKPLSPKMSLRNWNVWVTP SSPVPPTPASSTKISSPKALIPVSKSSQADACSESRNRKRLDSCLESVKQKC VSKCNVTDLQAEISRLDLCLSGTQVLSQDSEGTYSKSTEGAGTISISEPPRPV TPYASGGCPLPLPPLPCGSEMGKENSPEENKNWLLAIARAKRSESSSPSPSSQ TPSSRSQSGKTSPPVITTPSSMRKICTYFRKRKTQDDFCSPHSTEL"
Query Match	63.7%; Score 1802.2; DB 6; Length 2350;
Best Local Similarity	86.9%; Pred. No. 0;
Matches 2005; Conservative	0; Mismatches 298; Indels 4; Gaps 2;
QY	23 GATAACGATTGTGTGTGAGAGGCGCAACGTCGATTTCTGCTGAACCTTGAGGC-ATT 81
DB	3 GAAACCGCCCGGTTGTGGAGGTCACAGCGCGCTCTGTGGAACCTTGGGGCGATC 62
QY	82 TCTACGACTTTTCTCTCAGCTGAGGCTTTCTTCGACCCCTGTGCTCTTCAATTCGGTG 141
DB	63 TTCTCGCGCTTCCCTTCAGCTGAGCCTTCTCTCAGCCGACAGTGTCTTCAACTCGGTA 122
QY	142 CTCGCGCAGCCGCGCTTGGCGTCTTGAGAAATGGATGGTCTTCACAATACCCCTCTCAA 201
DB	123 CTCGCGCAGCCGCGCTTGGCGTCTTGAGAAACGGGTGTCTTCACTTACCCCTTACAA 182
QY	202 TCCCTTCTGACTGGTTATCAGTGCAGTGTGTAATGATGAACACACACTTCTTATGGAGAAACA 261
DB	183 TCCCTTCTGAGTGTATCAGTGCACCTGTACGATGAACACACAGCTCTTATGGAGAAACA 242
QY	262 GGAGTCCCAAGTCTCTCTTTGGATGATACCTTCTTCTGCTCCCAATATGGAAACATGTA 321
DB	243 GGAGTCCCAAGTCTCTCTTTGGATGATACCTTCTGACCGCTCCCAAGTATGGAGCATATA 302
QY	322 CTAGCAGTTGCCAATGAAGAAGCTTTGTTCAGTGTATTAACACAGAAATCACAAAGTTTC 381
DB	303 TTAGCAGTTGTCTAATGAAGAAGCTTTGTGAGATATATAATACAGAAATCACAAATCAGC 362
QY	382 AGAAGAAAGTGTCTCAAGAATGGATGGCTCACTTGAATGCCGTCTTTGACCTGGCGCTGG 441
DB	363 AAAAGACATGTTCAAGGAGTGGATGGCTCACTGGAATGCTGTCTTTGACTTGGCGTGG 422
QY	442 GTTCTGTGTAACCTTAAACTTTGTTACAGCAGCAGGTGATCAACAGCCAAAATTTTGGGAC 501
DB	423 GTCCCTGTGTAACTTAAACTTTGTTACAGCAGCAGCGGTGATCAGACAGCCAAAATTTGGGAT 482
QY	502 GTAAAGCTGTGTGAGTGTGAACATCAAGGTCAATGAGCCTCAAGTCAGT 561
DB	483 GTAAGAGCTGTGTGAGTGTGAACATCAAGGTCAATGAGCCTCAAGTCAGT 542
QY	562 GCCTTTCTTAAGTTTGAAGAAAGCTGTATTCTGTACGGGTGGAAGAGATGGCAACATTATG 621
DB	543 GCCTTTCCCAAGTTTCAAAAAGCTGTGTTCTCTACAGGGGGGAGAGACGGCAACATTATG 602
QY	622 GTCTGGATACCAAGGTGCAACAAAAAGATGGTTTTATAGGCAAGTGAATCAAAATCAGT 681
DB	603 ATCTGGGACACACAGGTGTAAACAAAAAGATGGATTTTATAGACAGTGAATCAAAATCAGT 662
QY	682 GGAGCTCACAATACCTCAGACAGCAAGCAACCCCTTCAAAACCCAGAGAGAAACAGATTCA 741
DB	663 GGAGCTCACAATACCTCAGACAGCAAGCAACCCCTTCAAAAGCCCAAGAGAAACAAATTTCA 722
QY	742 AAAGGACTTGTCTCTGCTGGATTTCACAGCAAGTGTGTGCTCTCTTTTCAAGAC 801
DB	723 AAAGGACTTGTCTCTGCTGGATTTCACAGCAGAGGTGTACTGTGCTCTCTTTTCAAGAT 782
QY	802 GAGAAATACCTTAGTCTCAGCAGGAGTGTGGATGGGATAATCAAGATATGGGATTTACGT 861
DB	783 GAGAAATACCTTAGTCTCAGCAGGAGTGTGGATGGGATAATCAAGATATGGGATTTGCGC 842
QY	862 AGAATATTACTGCTTATCGAAGAACCCATAGATCCAGTCTTCTCTGTACCCAGT 921
DB	843 AAGAAATTACACTGCTTATCGAAGAACCCATAGATCCAGTCTTCTCTGTACCCAGT 902
QY	922 AGCAGCACTCGAATAAATTGGATATTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 981

|||||
903 ACCAGACTCGAAAGCTAGGATCTCGAGTTGGTTTATAGACTCTACTGGCTCTACTTTA 962
982 TTTTGCTAAATTGCACAGACCAATCACTACATGTTTAAATATGACTGGGTGAAGACTTCT 1041
963 TTTTGCTAACTGCACAGATGACAACATCTATATGTTCAATATGACTGGCTTAAAGACTTCT 1022
1042 CCAGTGGCTATTTTCAATGGACACGAACTCTACTCTTTTATGTAATAATCCAGCCTTAGT 1101
1023 CCGGTGGCTGTCTTCAATGGACACCAAGAACTCTACCTTTTATGTAATAATCAAGTCTTAGT 1082
1102 CCAGATGACCACTTTTATGCTAGTGGCTCAAGTGATGAAGCTGCTACATATGGAAGGTC 1161
1083 CCAGATGACCACTTTTATGCTAGTGGCTCAAGTGATGAAGCTGCTACATTTGGAAGGTT 1142
1162 TCCACACCCCTGGCAACCTCTACTGTGCTCTGGGTCAATTTCAAGAGAGTCAAGTCTGTG 1221
1143 TCCATGCCATGGATCTCTACTGTGCTCTGGGTCAATTTCAAGAGGTCAAGTCTGTG 1202
1222 TGCTGGTGCATCTGATCTTCAAGAGATGCTACTGTCTGTGATGACAAATACACTAA 1281
1203 TGCTGGTGCATCTGATCTTCAAGAGATGCTACTGTCTGTGATGACAAATACACTGAAA 1262
1282 ATCTGGCGCTTCAATAGAGCTTAGAGGAGAAACAGGAGGTGATAAATTTCCACGGTG 1341
1263 ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAAC---AGGTGATAAAATTTCCATAGTG 1319
1342 GGTGGGCTCTCAGAAAGAAAAGAGTCAAGACCTGGCTAGTAAACAGTACAGAGTAGC 1401
1320 GGTGGACCTCTCAGNAGAAAAGAGTGAAGGCTGCCAGTAAAGTACCAAGTAGC 1379
1402 CAGAGTACTCTGCCAAAGCCCGAGGTAAGTGAATGCCATCCAAATTTCTTCCCGGTCA 1461
1380 CAGAGTACTCTGCCAAAGCTCCAGAGCAAGAGCAGTCCATCCATCTCTCTCTCTCG 1439
1462 TCCGAGCTTGTGCCCCAGAGCTGTGCTGAGACCTCCCTCTCTTCCATTAATCTCTAGG 1521
1440 TCAGCAGCTTGTACTCCGAGCTGTGAGAGACCTCCCTCTTCCATTAAGTACCCCA 1499
1522 TTCTCTATTAAACCTCTCTGCGAAGGCCCGGTCTCCCATCAACAGAGAGGCTCTGTG 1581
1500 TTCTCAGTCAAAACCACTCTGCCAGCACCCTTCTTCTCAGTCAGCAGAGAGGCTCCATC 1559
1582 TCTCCTCGTCTCTCCAAAGCCACTTCAATCTTTCAAGATGTCGATTAGAAATCGGTGACC 1641
1560 TCTTCTGTCTCTCCAAAGCACTCTCATCTTTCAAGATGTCGCTTAGAAATCGGTGACC 1619
1642 CGAAGCACTTCTCATCACCACCCATCACTCCACCTGCTTCGAGACCAAGATCATGCT 1701
1620 CGAAGCACTTCTCTCATCACCACCTGTCCTCCTGCTTCTGAGCAAAAGATCTCATCT 1679
1702 CCGAGAAAAGCCCTTATCTCTGTGACCCAGAGTCACTCCCAAGCAGAGGCTTGTCTGTAG 1761
1680 CCAAGAAAAGCTTATTCTGTGACCCAGAGTCACTACAGCAGATGCTTGTCTGTAA 1739
1762 TCTAGAAATAGAGTAAAGAGGCTTAGACTCAAGCTGTCTGAGAGTGTGAAAACAAAG 1821
1740 TCTAGAAATAGAGTAAAGAGGCTTGTACTCAAGCTGTCTGAGAGTGTGAAAACAAAG 1799
1822 TGTGTGAAGAGTTGTAACCTGTGTGACTGAGCTGTGAGCCAGTTGAAAATCTTCAATTG 1881
1800 TGTGTGAAGAGTTGCAACTGTGTACTGTGCTTGTGACCGCAAGCGGAGAGTCTTGTG 1859
1882 GATCTGTGCTGCTCTGCTGTAACCAAGGAGACCTTAGTAAGGACTCTCTAGTCTCTACC 1941
1860 GATCTGTGCTGCTTCTCCGCAACCCAGAGAGCTCTTAGCCAAGACTCCGAGGCTCTACC 1919
1942 AAATCAAGCAAAATTAAGAGAGCTGTAACAGTATCTCAGAGCCTCCGCTCTCTATCAGT 2001
1920 AAATCAAGCAAGACTGAAGGTGTGGCACAAGCATCTCAGAACCTCTCTCTCTGTGAGT 1979
2002 CCGTATGCTTCAGAAAGCTGTGAACGCTACCTCTCTTCTTGTGAGACCTTGTGGAGAGGG 2061
|||||

Db 1980 CCTATGCTTCTGAAGCTGTGGACCACTGCCTCTTCTCTTGTGAGACCTTGTGGAGAGGA 2039
Qy 2062 TCTGAATGTGTAGGCAAGAAATAGTTTCCCAGAGAAATAAAACTGTTGTGGCCATG 2121
Db 2040 TCTGAGATGTGTGGGCAAGAAATAGCTTCTCAGAGAAATAAGAACTGTTGTGGCCATA 2099
Qy 2122 GCAGCCAAACGGAAGGCTGGAATCCATCTCCAGAGTCCGTCATCCAGACACCCCAAT 2181
Db 2100 GCAGCCAAACGGAAGGCAAGAAATTTTATCCCCAAGAGTCCATCATCTCAGACACCCAGT 2159
Qy 2182 TCCAGGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTCAACATCCAGCCAGCTCCATG 2241
Db 2160 TCCAGGAGACAAAGTGGGAAGACGTCACAGGCCCGGTCAACATTTACTTCTCCAGCTCCATG 2219
Qy 2242 AGGAAATCTGCATACATCTTCCATAGAAAAGTCCAGGAGGACTTCTGTGGTCTCTGAACAC 2301
Db 2220 AGGAAGATATGTACATATCTTTCGTAGAAAGACTCAAGATGACTTCTGCACTCTCTGAACAC 2279
Qy 2302 TCACAGAAATATAGATTCTTAATCTGA 2328
Db 2280 TCAACTGAATATTAGATGCTAATCTGA 2306

RESULT 4
AK164249 4179 bp mRNA linear HTC 21-SEP-2005
DEFINITION Mus musculus 9 days embryo whole body cDNA, RIKEN full-length enriched library, clone:D030041C04 product:l2DPL protein, full insert sequence.

ACCESSION AK164249
VERSION AK164249.1 GI:74211282
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
PUBMED 10349636

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
PUBMED 11042159

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Tozawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
PUBMED 11076861

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yananaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, M., Gaasterland, T., Glissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Mateu, Y., Nikaudo, I., Pesole, G., Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,

Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kaniya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P., Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H., Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H., Toyooka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S. and Hayashizaki,Y.

RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium

Functional annotation of a full-length mouse cDNA collection
Nature 409 (5821), 685-690 (2001)
11217851

5

Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S., Kikado,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I., Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A., Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C., Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H., Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V., Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A., Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T., Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S., Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A., Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A., Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R., Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T., Numata,K., Okido,T., Pavan,W.J., Pertea,G., Pesole,G., Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S., Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M., Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K., Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,K., Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y., Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I., Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P., Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M., Sakazume,N., Sato,K., Shiraka,T., Waki,K., Kawai,J., Aizawa,K., Akawa,T., Fukuda,I., Hazara,A., Hashizume,W., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Miyazawa,A., Sakai,K., Sasaki,D., Shibata,K., Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S., Rogers,J., Birney,E. and Hayashizaki,Y.

FANTOM Consortium

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420 (6915), 563-573 (2002)
12466851

6

Carninci,P., Kasukawa,T., Katayama,S., Gough,J., Frith,M.C., Maeda,N., Oyama,R., Ravasi,T., Lenhard,B., Wells,C., Kodzius,R., Shimokawa,K., Bajic,V.B., Brenner,S.E., Batalov,S., Forrest,A.R., Zavolan,M., Davis,M.J., Wilming,L.G., Aidinis,V., Allen,J.E., Ambesi-Impombato,A., Apweiler,R., Aturaliya,R.N., Bailey,T.L., Bansal,M., Baxter,L., Beisel,K.W., Bersano,T., Bono,H., Chalk,A.M., Chiu,K.P., Choudhury,V., Christoffels,A., Clutterbuck,D.R., Crowe,M.B., Dalla,E., Dalrymple,B.P., de Bono,B., Della Gatta,G., di Bernardo,D., Down,T., Engstrom,P., Fagioli,M., Faulkner,G., Fletcher,C.F., Fukushima,T., Furuno,M., Futaki,S., Gariboldi,M., Georgii-Hemming,P., Gingeras,T.R., Gojobori,T., Green,R.E., Gustincich,S., Harbers,M., Hayashi,Y., Hensch,T.K., Hirokawa,N., Hill,D., Humintick,L., Iacono,M., Ikeo,K., Iwama,A., Ishikawa,T., Jakt,M., Kanapin,A., Katoh,M., Kawasawa,Y., Kelso,J., Kitamura,H., Kitano,H., Kollias,G., Krishnan,S.P., Kruger,A., Kummerfeld,S.K., Kurochkin,I.V., Lareau,L.F., Lazarevic,D., Lipovich,L., Liu,J., Liuni,S., McWilliam,S., Madan,Babu,M., Madera,M., Marchionni,L., Matsuda,H., Matsuura,S., Miki,H., Mignone,F., Miyake,S., Morris,K., Motagui-Tabar,S., Mulder,N., Nakano,N., Nakachi,H., Ng,P., Nilsson,R., Nishiguchi,S., Nishikawa,S., Nori,F., Ohara,O., Okazaki,Y., Orlando,V., Pang,K.C., Pavan,W.J., Pavoni,G., Pesole,G., Petrovsky,N., Piazza,S., Reed,J., Reid,J.F., Ring,B.Z., Ringwald,M., Rost,B., Ruan,Y., Salzberg,S.L., Sandelin,A., Seno,S., Seesla,I., Sheng,Y., Shibata,Y., Shimada,H., Shimada,K., Silva,D., Sinclair,B., Sperling,S., Stupka,E., Sugiuira,K., Sultana,R.,

Db 1 GAGTTGGAGCGAAACCGCGCGGTTGTGGAGGGTGACAGCGCGGCTCTCTGTGGAACT 60
QY 72 TGGAGGC-ATTTCTAGGACTTTTCTCTCAGCTGAGGCTTTTCTCGGACCCCTGATGCTCT 130
Db 61 TGGGGGCGAGTCTCTCGGCGGTTCCCTCAGCTGAGCGCTTCTCCTCAGGCCCCAGATGCTCT 120
QY 131 TCAATTCGGTGCTCGCGCAGCCCGAGCTTGGCGTCTCTGAGAAATGGATGGTCTTCCACAAT 190
Db 121 TCAACTCGGTACTTCGCGCAGCGCAGCTCGGCGTCTCTGAGGAAACGGGTGGTCTTTCACATT 180
QY 191 ACCCTCTTCAATCCCTTCTGACTGGTATCAGTCAGATGGTAAATGATGAACACACTTCTT 250
Db 181 ACCCTCTCAATCCCTTCTTAAGTGGTTATCAGTGGCAACTGTAACGATGAACACACAGTCTT 240
QY 251 ATGGAGAAACAGGAGTCCAGTTCTCTCTTTTGGATGTACCTTCTCTTCTGCTCCCAATA 310
Db 241 ATGGAGAAACAGGAGTCCAGTCTCTCTTTTGGATGCACCTTCTGTACCGCTCCCGAT 300
QY 311 TGGAAACATGTACTAGCAGTTGCCAATGAAGAAGGCTTTGTTTCGATTGTATACACAGAAT 370
Db 301 TGGAGCATATATTAGCAGTTGCTAATGAAGAAGGCTTTGTTCAGATTATATAATACAGAAT 360
QY 371 CACAAAGTTTCAGAAAGAGTCTTCAAGAAATGGATGGCTCACCTGGAATGCCGCTTTTG 430
Db 361 CACAAACTAGCAAAAGACATGCTTCAAGGAGTGGATGGCTCACTGGAATGCTGCTTTG 420
QY 431 ACCTGGCTGGGTCTCTGTGAACTTTAACTTTGTTACAGCAGCAGGTGATCAAAACAGCA 490
Db 421 ACTTGGCTGGGTCTCTGTGAACTTTAACTTTGTTACAGCAGCAGGTGATCAACAGCA 480
QY 491 AATTTTGGACGTAAAGCTGTGTAGCTGATTGGAAACATGCAAAAGTCAATCAATCGAGC 550
Db 481 AATTTTGGATGTAGAGCTGTGTAGCTGATGGGACATGCAAAAGCCACCAAGTCAAGC 540
QY 551 TCAA-GTCAGTTGCCCTTTTCTAAGTTTGAAGAAGCTGTATTCTGTACGGGTGGAAGAGAT 609
Db 541 TCAAGTCTGTAGCTTTTCCCGAGTTTCAAAAGCTGTGTTCTCTACAGGGGGAGAGAC 600
QY 610 GGCACATTATGGTCTGGGATACGAGTGCACAAAGGATGGGTTTTATAGGCAAGTG 669
Db 601 GGCACATTATGATCTGGGACACGAGTGTAAACAAAAGAGTGGATTTTATAGACAAAGT 660
QY 670 AATCAAACTCAGTGGAGCTCACAACTCTCAGACAAAGCAACCCCTTCAAAACCCCAAGAG 729
Db 661 AATCAAACTCAGTGGAGCTCACAACTCTCAGACAAAGCAACCCCTTCAAAAGCCCAAGAG 720
QY 730 AAACAGAAATTCAAAAGGACTTGTCTCTCTGTGGATTTTCAGCAAAAGTGTACTGTGGTC 789
Db 721 AAACAAAATTCAAAAGGACTTGTCTCTCTGTGGATTTCCACAGAGTGTACTGTGGTC 780
QY 790 CTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGAGAGCTGTGGATTAATCAAGTA 849
Db 781 CTCTTTCAAGATGAGAATACATTAGTCTCAGCAGGAGCGGTGGATGGAATATCAAAAGTA 840
QY 850 TGGGATTTTACGTAAGAAATATATCTGCTTATCCAGAGAACCCATAGCATCCCAAGTCTTTC 909
Db 841 TGGGATTTTGGCAAGAAATATACACTGCTTATCAGCAGAAACCCATAGCATCCCAAGTCTTTC 900
QY 910 CTGTACCCAGGTAGCAGCACTCGAAAACCTTGGATATTTCAAGTCTGATTTTGGATTCCACT 969
Db 901 CTGTACCCAGGTATCCAGCACTCGAAAGCTAGGATCTCGAGTTGGTTTTAGACTCTACT 960
QY 970 GGCCTACTTTTATTTGCTAAATTTGCACAGACGATACATCTACATGTTTAAATATGACTGG 1029
Db 961 GGCCTACTTTTATTTGCTAACTGCACAGATGCAACATCTATATGTTTCAATATGACTGGC 1020
QY 1030 TTGAAGACTTCTCCAGTGGCTATTTTCAATGACACACGAACTCTACCTTTTATGTAAAA 1089
Db 1021 TTAAAGACTTCTCCGGTGGCTGTCTCAATGACACACGAACTCTACCTTTTATGTAAAA 1080
QY 1090 TCCAGCTTTAGTCCAGATACCAAGTTTATTTAGTCAGTGGCTCAAGTGAAGCTCCCTAC 1149
Db 1081 TCAAGTCTTTAGTCCAGATACCAAGTTTATTTAATCAGTGGTTTCAAGTGTGAAGCTGCCTAC 1140

QY 1150 ATATGGAGGTTCTCACACCTGSCAACCTCTCTACTGTGCTCTGGGTCAATTTCTCAAGAG 1209
Db 1141 ATTTGGAAAGTTTTCATGCCATGGATCTCTCTACTGTGCTCTGGGTCAATTTCTCAAGAG 1200
QY 1210 GTCACGTCTGTGTGCTGGTGTCCATCTGACTTTCAAAAGATTGCTACCTGTTCTGTGATGAC 1269
Db 1201 GTCACGTCTGTGTGCTGGTGTCCATCAGACTTTCACCAAGATTGCAACCTGTCTGTGATGAT 1260
QY 1270 AATACACTAAAAATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACCAAGGAGGTGATAAA 1329
Db 1261 AATACACTGAAAAATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACC--AGGTGATAAA 1317
QY 1330 CTTTCCACGGTGGGTGGGCTCTCAGAGAAAGAAAGAGTCAAGACCTGGCTTACTATAA 1389
Db 1318 CATTTCCATAGTGGGTGGACCTCTCAGAAAGAAAGAAAGTGAAGACCTGCCAGTAAG 1377
QY 1390 GTAACGAGTAGCCAGATPACTCTCTGCCAAAGCCCCCAGGGGTAAAGTGCATTCCTCAAT 1449
Db 1378 GTACCAAGTAGCCAGAGTACTCTCTGCCAAAGCTCCCAGAGCCNAGAGCAGTCCATCCATC 1437
QY 1450 TCTTCCCCGTCAATCCGACGCTTGTGCCCAAGCTGTGTCTGAGACCTCCCTCTTCTCTTCA 1509
Db 1438 TCTCTCTCTCGTCAGCAGCTTGTACTCCGAGCTGTGCAGGAGACCTCCCTCTTCTCTTCA 1497
QY 1510 AATACTCTTACGTTTCTCTATTAAACCTCTCTCTGCCAAGGCCCGTCTCCCATCAACAGA 1569
Db 1498 AGTACCCCCCAATTCTCAGTCAAAACCACTCTCTGCCAAGCCACTCTCATCTTCAAGATGTGCTTAGA 1617
QY 1570 AGAGCTCTGTCTCTCTCCGCTCTCTCCAAAGCCACCTTCTCATCTTTCAAGATGTGCAATTAGA 1629
Db 1558 AGAGCTCCATCTCTTCTGTGTCTCTCCAAAGCCACTCTCATCTTCAAGATGTGCTTAGA 1617
QY 1630 AACTGGGTGACCCGAAACACTTCTCTCATCACACCCATCTCACCTGCTTGGAGACC 1689
Db 1618 AACTGGGTGACCCGAAACACTTCTCTCATCACACCTGTCTCACTCCACCTGCTTCTGAGACA 1677
QY 1690 AAGATCATGTCTCCGAGAAAGCCCTTATTCTCTGTGAGCCAGAGTCAATCCCAAGCAGAG 1749
Db 1678 AAGATCTCATCTCCAAGAAAGCTCTTATTCTGTGAGCCAGAGTCAATCAGGACAGAT 1737
QY 1750 GCTTGTCTCTGAGTCTAGAAATAGAGTAAAGAGGAGGCTTAGACTCAAGCTGTCTGGAGAGT 1809
Db 1738 GCTTGTCTCTGAATCTAGAAATAGAGTGAAGAGGCTCTTGACTCAAGCTGTCTGGAGAGT 1797
QY 1810 GTGAAAACAAAAGTGTGAAAGATTGTAACCTGTGTGACTGAGCTTGTGAGCCCAAGTTGAA 1869
Db 1798 GTGAAAACAAAAGTGTGAAAGATTGCAACTGTGTCTCACTGAGCTTGTGACGCCCAAGCGAG 1857
QY 1870 AATCTTCATTTGGATCTGTGCTGCTTGTCTGTAAACCAAGGAAGACCTTAGTAAGGACTCT 1929
Db 1858 AGTCTTCTGTTTGGATCTGTGCTGCTTTTCCGGCACCCAGGAAGTCTCTTAGCCAAAGCTCC 1917
QY 1930 CTAGTCTCTACAAATCAAGCAAAATTTGAAGGAGCTGGTACCAGTATCTCAGAGCCTCCG 1989
Db 1918 GAGGTCCTACCAATCAAGCAAGCTGAGGTGTGSCACAGCATCTCAGAACTCTCT 1977
QY 1990 TCTCTATCAGTTCGGTATGCTTTCAGAAAGCTGTGGAACGCTACCTCTTCTCTTTGAGACT 2049
Db 1978 TCTCTGTCACTCTTATGCTTTCGAAGGCTGTGACCACTGCCTCTTCTCTTTGAGACT 2037
QY 2050 TGTGAGAGAGGCTCTGAAATGTTAGGCAAGAGATAGTCTCCCAAGAGAAATAAAACCTGG 2109
Db 2038 TGTGAGAGAGGATCTGAGATGTTGGGCAAGAGAAATAGCTCTTCCAGAGAAATAGAACTGG 2097
QY 2110 TTGTTGGCCATAGGAGCCAAACGGAAGCTGAGAAATCCATCTCCACGAAGTCCGCTCATCC 2169
Db 2098 TTGTTGGCCATAGCAGCCAAACGCAAGGAGCAAAATTCATCCCAAGAGTCCATCATCT 2157
QY 2170 CAGACACCAATTTCCAGAGACAGAGCGGAAAGACATTTGCCAAGCCCGTCCACCATCAG 2229
Db 2158 CAGACACCCAGTTCCAGGAGCAAAAGTGGGAAGAGCTCACCGAGGCCCGGTCCACCTACT 2217

QY	2230	CCGAGCTCCATGAGGAAATCTGCACATACCTTCATAGAAAGTCCAGGAGGACTTCTGT	2289
Db	2218	CCAGCTCCATGAGGAGATATGTACATCTTCGTAGAAAGACTCAAGTACACTTCTGC	2277
QY	2290	GGTCCTGACACTCAACAGAAATATAGATTCTTAATCTGA	2328
Db	2278	AGTCCTGAACACTCAACTGAATATATAGATGCTAATCTGA	2316
RESULT 5			
AKI69601			
LOCUS	AKI69601	4201 bp	linear
DEFINITION	Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN full-length enriched library, clone:E430001J12 product:L2DTL protein, full insert sequence.		
ACCESSION	AKI69601		
VERSION	AKI69601.1	GI:74143874	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Murioidea; Murinae; Mus.		
REFERENCE	1		
AUTHORS	Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
PUBLISHED	10349636		
REFERENCE	2		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
PUBLISHED	11042159		
REFERENCE	3		
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tachiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Ogawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, K., Ishikawa, T., Tanaka, T., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
PUBLISHED	11076861		
REFERENCE	4		
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Harai, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fletschmann, W., Gaasterland, T., Giesi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J.S., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Botunga, N., Carninci, P., de Bona, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, C., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohtsuki, S. and Hayashizaki, Y.		
CONSORTIUM	RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium		
TITLE	Functional annotation of a full-length mouse cDNA collection		
JOURNAL	Nature 409 (6821), 685-690 (2001)		
PUBLISHED	11217851		
REFERENCE	5		

AUTHORS	Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Sempie, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, K., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Harai, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
CONSORTIUM	FANTOM Consortium
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420 (6915), 563-573 (2002)
PUBLISHED	12466851
REFERENCE	6
AUTHORS	Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shiolkawa, K., Bajic, V.B., Brenner, S.E., Batalov, S., Forrest, A.R., Zavolan, M., Davis, M.J., Wilming, L.G., Aidinis, V., Allen, J.E., Ambesi-Impombato, A., Apweiler, R., Aturaliya, R.N., Bailey, T.L., Banal, M., Baxter, L., Beisel, K.W., Bersano, T., Bono, H., Chalk, A.M., Chiu, K.P., Choudhary, V., Christoffels, A., Clutterbuck, D.R., Cline, M.P., Dall'Aglio, E., Dalrymple, B.P., de Bono, B., Della Gatta, G., Di Bernardo, D., Down, T., Engstrom, P., Fagiolini, M., Faulkner, G., Fletcher, C.F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgii-Hemming, P., Gingeras, T.R., Gojobori, T., Green, R.E., Gustincich, S., Harbers, M., Hayashi, Y., Hensch, T.K., Hirokawa, N., Hill, D., Huminecki, L., Iacono, M., Ikeo, K., Iwama, A., Ishikawa, T., Jakt, M., Kanapin, A., Katoh, M., Kawasawa, Y., Keiso, J., Kitamura, H., Kitano, H., Kollias, G., Krishnan, S.P., Kruger, A., Kummerfeld, J., Kurochkin, I.V., Lareau, L.F., Lazarevic, D., Lipovich, L., Liu, J., Liuni, S., McWilliam, S., Madan, Babu, M., Madera, M., Marchionni, L., Matsuda, H., Matsuzawa, S., Miki, H., Mignone, F., Miyake, S., Moris, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakachi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Pang, K.C., Pavan, W.J., Pavese, G., Pesole, G., Petrovsky, N., Piazza, S., Reed, J., Reid, J.F., Ring, B.Z., Ringwald, M., Rost, B., Ruan, Y., Salzberg, S.L., Sandelin, A., Schneider, C., Schonbach, C., Sekiguchi, K., Sempie, C.A., Seno, S., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sinclair, B., Sperling, S., Stupka, E., Sugitara, K., Sultana, R., Takenaka, Y., Taki, K., Tammoja, K., Tan, S.L., Tang, S., Taylor, M.S., Tegner, J., Teichmann, S.A., Ueda, H.R., van Nimwegen, E., Verardo, R., Wei, C.L., Yagi, K., Yamanishi, H., Zabarovsky, E., Zhu, S., Zimmer, A., Hilde, W., Bult, C., Grimmond, S.M., Teasdale, R.D., Liu, E.T., Brusic, V., Quackenbush, J., Wahlestedt, C., Mattick, J.S., Hume, D.A., Kai, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori-Katayama, M., Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kato, T., Kawaji, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondo, S., Konno, H., Nakano, K., Nimmiya, N., Nishio, T., Okada, M., Plessy, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Watahiki, A., Okamura-Ohno, Y., Suzuki, H., Kawai, J. and Hayashizaki, Y.
CONSORTIUM	FANTOM Consortium
TITLE	The transcriptional landscape of the mammalian genome
JOURNAL	Science 309 (5740), 1559-1563 (2005)

16141072	251	ATGGAGAAACAGAGTCCAGTTCCTCTCTTTTGGATGTACCTTCTCTTCTGCTCCCAATA	310
PUBMED	QY		
REFERENCE			
AUTHORS			
Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C.C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T., Pang, K.C., Hallinan, J., Mattick, J., Hume, D.A., Lipovich, L., Batalov, S., Engstrom, P.G., Mizuno, Y., Faghihi, M.A., Sandelin, A., Chalk, A.M., Mottagui-Tabar, S., Liang, Z., Lenhard, B. and Wahlestedt, C.	Db		
CONSRMTM	QY		
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hori, F., Iida, J., Imamura, K., Imotani, K., Itoh, M., Kanagawa, S., Kawai, J., Kohjima, M., Konno, H., Murata, M., Nakamura, M., Ninomiya, N., Nishiyori, H., Nomura, K., Ohno, M., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shiraki, T., Tagami, M., Tagami, Y., Waki, K., Wataniki, A., Muramatsu, M. and Hayashizaki, Y.	Db		
TITLE	QY		
JOURNAL			
COMMENT			
source			
FEATURES			
misc_feature			
ORIGIN			
Query Match	62.8%	Score 1777.8;	DB 6; Length 4201;
Best Local Similarity	86.0%	Pred. No. 0;	
Matches 2016;	Conservative	0; Mismatches 297;	Indels 30; Gaps 3;
QY	12	GAGTTGAGGCGATAACGATTGTTGTGTGAGAGCGCAACGTCGGATTCTGCTGAAC	71
Db	1	GAGTTGAGGCGAAACCGCGCGCGTGTGTGAGGCGTGCAGACGCGCGCTCTCTGTGAACT	60
QY	72	TGGAGGC-ATTCTACGACTTTTCTTCAGCTGAGGCTTTTCTCCGACCTCATGCTCT	130
Db	61	TGGGGGCGGATCTTCGCGGCTTCCCTCAGCTGAGCGCTTCTCTCAGGCCACGATGCTCT	120
QY	131	TCAATTGCGTCTCCGCGAGCCGAGCTTGGCGTCCCTGAGAAATGGATGCTCTTCAAA	190
Db	121	TCAACTCGTCTCCGCGAGCCGAGCTCGCGTCTCGAGAACGGGTGCTTCAATT	180
QY	191	ACCCTCTTCAATCCCTTCTGACTGGTTATCAGTCAGTCAGTGGTAAATGATGAACACACTTCTT	250
Db	181	ACCCTCTCAAAATCCCTTCTAAGTGGTTATCAGTCAACTGTGAACGATGAACACACGCTCTT	240

Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kwasawa, Y., Kedierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Perte, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Read, J.C., Read, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, K., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

PANTOM Consortium
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420 (6915), 563-573 (2002)
12466851

6
Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., Kodzius, R., Shinkawa, K., Bajic, V.B., Brenner, S.E., Batalov, S., Forrest, A.R., Zavolan, M., Davis, M.J., Wilming, L.G., Aidinis, V., Allen, J.E., Ambesi-Imbombo, A., Apweiler, R., Aturaliya, R.N., Bailey, T.L., Banerji, M., Baxter, L., Beisel, K.W., Bersano, T., Bono, H., Chalk, A.M., Chiu, K.P., Choudhary, V., Christoffels, A., Clutterbuck, D.R., Crowe, M.L., Dalla, E., Dalrymple, B.P., de Bono, B., Della Gatta, G., di Bernardo, D., Down, T., Engstrom, P., Fagiolini, M., Faulkner, G., Fletcher, C.F., Fukushima, T., Furuno, M., Futaki, S., Gariboldi, M., Georgii-Hemming, P., Gingeras, T.R., Gojobori, T., Green, R.E., Gustincich, S., Harbers, M., Hayashi, Y., Hensch, T.K., Hirokawa, N., Hill, D., Humnietchi, L., Iacono, M., Ikeo, K., Iwano, A., Ishikawa, T., Jakt, M., Kanapin, A., Katoh, M., Kawasawa, Y., Keiso, J., Kitamura, H., Kitano, H., Kollias, G., Krishnan, S.P., Kruger, A., Kummerfeld, S.K., Kurochkin, I.V., Lareau, L.F., Lazarevic, D., Lipovich, L., Liu, J., Liuni, S., McWilliam, S., Madan, Babu, M., Madera, M., Marchionni, L., Matsuoka, H., Matsuzawa, S., Miki, H., Mignone, F., Miyake, S., Moris, K., Mottagui-Tabar, S., Mulder, N., Nakano, N., Nakauchi, H., Ng, P., Nilsson, R., Nishiguchi, S., Nishikawa, S., Nori, F., Ohara, O., Okazaki, Y., Orlando, V., Pang, K.C., Pavan, W.J., Pavese, G., Pesole, G., Petrovsky, N., Piazza, S., Reed, J.F., Reid, J.F., Ring, B.Z., Ringwald, M., Rost, B., Ruan, Y., Salzberg, S.L., Sandelin, A., Schneider, C., Schonbach, C., Sekiguchi, K., Semple, C.A., Seno, S., Sessa, L., Sheng, Y., Shibata, Y., Shimada, H., Shimada, K., Silva, D., Sinclair, B., Sperling, S., Stupka, E., Sugita, K., Sultana, R., Takenaka, Y., Taki, K., Tammoja, K., Tan, S.L., Tang, S., Taylor, M.S., Tegner, J., Teichmann, S.A., Ueda, H.R., van Nimwegen, E., Verardo, R., Wei, C.L., Yagi, K., Yamanishi, H., Zabarovsky, E., Zhu, S., Zimmer, A., Hide, W., Bult, C., Grimmond, S.M., Teasdale, R.D., Liu, E.T., Brusic, V., Quackenbush, J., Wahlestedt, C., Mattick, J.S., Hume, D.A., Kai, C., Sasaki, D., Tomaru, Y., Fukuda, S., Kanamori-Katayama, M., Suzuki, M., Aoki, J., Arakawa, T., Iida, J., Imamura, K., Itoh, M., Kato, T., Kawaji, H., Kawagashira, N., Kawashima, T., Kojima, M., Kondo, S., Konno, H., Nakano, K., Ninomiya, N., Nishio, T., Okada, M., Plessey, C., Shibata, K., Shiraki, T., Suzuki, S., Tagami, M., Waki, K., Watahiki, A., Okamura-Oho, Y., Suzuki, H., Kawai, J. and Hayashizaki, Y.

FANTOM Consortium
The transcriptional landscape of the mammalian genome
Science 309 (5740), 1559-1563 (2005)
16141072

7
Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C.C., Suzuki, M., Kawai, J., Suzuki, H., Carninci, P., Hayashizaki, Y., Wells, C., Frith, M., Ravasi, T., Pang, K.C., Hallinan, J., Mattick, J., Hume, D.A., Lipovich, L., Batalov, S., Engstrom, P.G., Mizuno, Y., Faghihi, M.A., Sandelin, A., Chalk, A.M., Mottagui-Tabar, S., Liang, Z., Lenhard, B. and Wahlestedt, C.

CONSRMTM
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS

RIKEN Genome Exploration Research Group
Antisense transcription in the mammalian transcriptome
Science 309 (5740), 1564-1566 (2005)
16141073

8 (bases 1 to 4228)
Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hori, F., Iida, J., Imamura, K., Imotani, K., Itoh, M., Kanagawa, S., Kawai, J., Kojima, M., Konno, H., Murata, M., Nakamura, M., Ninomiya, N., Nishiyori, H., Nomura, K., Ohno, M., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shiraki, T., Tagami, M., Tagami, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission
Submitted (30-MAR-2004) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan (E-mail: genome-res@sc.riken.jp, URL: http://genome-gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site for further details.
URL: http://genome-gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.

Location/Qualifiers
1. 4228
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="FANTOM DB:F630050P04"
/db_xref="taxon:10090"
/clone="F630050P04"
/cell_type="NOD-derived CD11c +ve dendritic cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"

COMMENT

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site for further details.
URL: http://genome-gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.

Location/Qualifiers
1. 4228
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="FANTOM DB:F630050P04"
/db_xref="taxon:10090"
/clone="F630050P04"
/cell_type="NOD-derived CD11c +ve dendritic cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"

misc_feature
1. 4228
/note="L2DPL protein (UniProt|Q8BW38, evidence: FASTY, 98.5%ID, 100%length, match=2221)
putative"

FEATURES
source

ORIGIN

Query Match 62.1%; Score 1758; DB 6; Length 4228;
Best Local Similarity 85.9%; Pred. No. 0;
Matches 2023; Conservative 0; Mismatches 295; Indels 38; Gaps 5;

QY 8 GCGGGAGTTGGAGCGGATACGATTTGTGTGTGAGAGCGGCAACGTCGATTTCTGTG 67
6 GCGGGAGTTGGAGCGGAAACCGCGCGGCTTTGTGGAGGTTGCAGACGGCGCTCTGTGG 65

DB 68 AACTTGGAGGC-ATTTCTACGACTTTTCTCAGCTGAGGCTTTTCTCCGACCCCTGATG 126
66 AACTTGGGCGCGATCTTCTCGCGCGTTCCTCCTCAGCTGAGCCTTCTCCTCAGGCCAGATG 125

QY 127 CTCCTTCAATTCGGTCTCGCGCAGCCAGCTTGGCGTCTTGAGAAATGGATGGTCTTCA 186
126 CTCCTTCAATTCGGTCTCGCGCAGCCAGCTGCGCGTCTTGAGGAAAGGGTGGTCTTCA 185

QY 187 CAATACCCCTCTTCAATCCCTTCTGACTGGTTATCAGTGCAGTGGTAATGATGAACACACT 246
186 CATTACCCCTTACAAATCCCTTCTAAGTGGTTATCAGTGCACATGTAACGATGAACACAG 245

QY 247 TCTTATGAGAAACAGGATCCAGTTCCTCTTTGGATGTACCTTCTCTTCTGCTCC 306
246 TCTTATGAGAAACAGGATGCCAGTTCCTCTCTTTGGATGACCTTCTGTACCCCTCC 305

QY 307 AATATGGAAACATGTACTAGCAGTTGCCCAATGAAGAAGCGTTTGTTCGATTTGATTAACACA 366

CONSRMTM
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS

Db 306 AGTATGGAGCATATATTAGCAGTTGCTAAATGAAGAGGCTTTGTGTCAGATTATATAATACA 365
Qy 367 GAATCACAAGATTTTCAAGAAAGAGTGTCTTCAAGAAAGTGGATGGCTCACTGCAATGCGTC 426
Db 368 GAATCACAAGTCTAGCAAAAGAGACATGCTTCAAGAGTGGATGGCTCACTGCAATGCTGTC 425
Qy 427 TTTGACCTGGCCCTGGGTTCTGGTGAACTTTAAACTTGTGTTACAGCAGCAGGTGATCAAA 486
Db 428 TTTGACCTGGCCCTGGGTTCTGGTGAACTTTAAACTTGTGTTACAGCAGCAGGTGATCAGACA 485
Qy 487 GCCAAATTTTGGGACGTAAAGCTGGTGAAGTGAATGGAACATGCAAGGTCAATGTC 546
Db 488 GCCAAATTTTGGGATGTAAGAGCTGGTGAAGTGAATGGAACATGCAAGGTCAATGTC 545
Qy 547 AGCTCAAGTCAGTTGCTTTTCTAAGTTTGAAGAAAG----- 583
Db 548 AGCTCAAGTCAGTTGCTTTTCTAAGTTTGAAGAAAG----- 583
Qy 584 -----CTGATTTCTGACCGGTGGAAGAGATGGCAACATTTATGCTTGGGATACCA 634
Db 606 CTCAAAATATCTGTTCTCTACAGGGGGAGAGACGGCAACATTTATGATCTGGGACACCA 665
Qy 635 GGTGCAACAAAAGATGGGTTTATAGGCAAGTGAATCAAAATCAAGTGGAGCTCAAA 694
Db 666 GGTGTAACAAAAGATGGGTTTATAGGCAAGTGAATCAAAATCAAGTGGAGCTCAAA 725
Qy 695 CCTCAGACAGCAACCCCTTCAAAACCCAGAGAGCAAGATTCAAAGAGCTTGGTC 754
Db 726 CTGAGACAGCAACCCCTTCAAAACCCAGAGAGCAAGATTCAAAGAGCTTGGTC 785
Qy 755 CTTCTGTGGATTTCCAGCAAGTGTACTGTGGTCTCTTTCAAGAGCAGAAATACCTTAG 814
Db 786 CTGCTGTGGATTTCCAGCAGAGTGTACTGTGGTCTCTTTCAAGAGCAGAAATACCTTAG 845
Qy 815 TCTCAGCAGAGCTGTGGATGGGATTAATCAAGATAGGAGTTTACGTAAGAAATTAAGT 874
Db 846 TCTCAGCAGAGCTGTGGATGGGATTAATCAAGATAGGAGTTTACGTAAGAAATTAAGT 905
Qy 875 CTTATCGACAGAACCCATAGCATCCAAAGTCTTTCTGTATCCAGGTACGAGCAGCTCGAA 934
Db 906 CTTATCGACAGAACCCATAGCATCCAAAGTCTTTCTGTATCCAGGTACGAGCAGCTCGAA 965
Qy 935 AACTTGGATATTCAGATCTGATTTTGGATTTCCACTGGCTCTACTTATTTGCTAAATGCA 994
Db 966 AGCTAGGATACTCGAGTTTGGTTTGTAGACTCTACTGGCTCTACTTATTTGCTAACTGCA 1025
Qy 995 CAGACGATACATCTACATGTTTAAATATGACTGGGTTGAAGACTTCTCCAGTGGCTATT 1054
Db 1026 CAGATGACACATCTATGTTCAATATGACTGGCTTAAAGACTTCTCCGGTGGCTGTCT 1085
Qy 1055 TCAATGGACACCAAGTCTTACCTTTTATGTAATCCAGCTTATGTCAGATGACCACT 1114
Db 1086 TCAATGGACACCAAGTCTTACCTTTTATGTAATCCAGCTTATGTCAGATGACCACT 1145
Qy 1115 TTTTATGTCAGTGGCTCAAGTGAAGTGCCTTACATATGGAAGTCTCCACACCCCTGGC 1174
Db 1146 TTTTATGTCAGTGGCTCAAGTGAAGTGCCTTACATATGGAAGTCTCCACATGGC 1205
Qy 1175 AACCTCTACTGTGCTCTGGGTCATTCTCAAGAGGTCAAGTCTGTGTGCTGGTGTCCAT 1234
Db 1206 ATCTCTCTACTGTGCTCTGGGTCATTCTCAAGAGGTCAAGTCTGTGTGCTGGTGTCCAT 1265
Qy 1235 CTGACTTCAAAAGATGCTTACCTTTTATGTAATCCAGCTTATGTCAGATGACCACT 1294
Db 1266 CAGACTTCAAAAGATGCTTACCTTTTATGTAATCCAGCTTATGTCAGATGACCACT 1325
Qy 1295 ATAGAGCTTAGAGGAGAACCCAGGAGGTGATATAAATTTTCCAGGTTGGGCTCTC 1354
Db 1326 ATAGAGCTTAGAGGAGAACCC--AGGTGATATAATTCATGATGGGTTGAGCTCTC 1382
Qy 1355 AGAAGAAAAAG-AGTCAAGACCTGGCTAGTAACAGTACGAGTACGAGTACTCCT 1413
Db 1383 AGAAGAAAAAGTGAAGCTGCGCCAGCTAACGGTACCAAGTAGCCAGAGTACTCCT 1442

Qy 1414 GCCAAGCCCCCAGGGTAAAGTGAATCCATCCAATTTTCCCGCTCATCCCGAGCTTGT 1473
Db 1443 GCCAAGCTCCAGAGCCAGAGCAGTCCATCCATCTCTCTCTCTCGTGCAGAGCTTGT 1502
Qy 1474 GCCCAAGCTGTGTGAGAGACCTCCCTCTTCTTCAAAATACCTCCTACGTTCTCTATAAA 1533
Db 1503 ACTCCAGCTGTGAGAGACCTCCCTCTTCTTCAAGTACCCCAATCTCTCAGTCAAA 1562
Qy 1534 ACTCTCTCTGCCAAGCCCGGCTCCCATCAACAGAGAGGCTCTGTCTCTCCGCTCT 1593
Db 1563 AACCTCTCTGCCAAGCCCGGCTTCTCAGTCAAGAGAGGCTCCATCTCTCTCTGTGTCT 1622
Qy 1594 CCCAAGCCACTTCTATCTTTCAAGATGTCCATTTAGAAATGGGTGACCCGGAACACCTTCC 1653
Db 1623 CCCAAGCCACTTCTATCTTTCAAGATGTCCATTTAGAAATGGGTGACCCGGAACACCTTCC 1682
Qy 1654 TCATCAACCACTCATCTCCACTGCTTCCGAGACCAAGATCATGTCTCCGAGAAAAGCC 1713
Db 1683 TCATCAACCACTGTCTCATCTCCACTGCTTCTGAGACCAAGATCTCATCTCCAAGAAAAGCT 1742
Qy 1714 CTTATTTCTGTGAGCCAGAACTCATCCACTGCTTCCGAGACCAAGATCATGTCTCCGAGAAAAGCC 1773
Db 1743 CTTATTTCTGTGAGCCAGAACTCATCCAGGAGATGCTTGTCTCTGAATCTAGAAATAGA 1802
Qy 1774 GTAAAGAGGAGCTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAGTGTGTGAAGAGT 1833
Db 1803 GTGAAGAGGCTGTGACTCAAGCTGTCTGGAGAGTGTGAAACAAAGTGTGTGAAGAGT 1862
Qy 1834 TGTAACTGTGTGACTGAGCTTGAATGAGCTTGAATGAGAAATCTTCAATTTGAGTCTGTGCTGC 1893
Db 1863 TGCACCTGTGTCACTGAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 1922
Qy 1894 CTTGTGTGTAAACAGGAGACCTTAG- TAAGAGCTCTCTAGGCTCTTACCAATCAAGCAA 1952
Db 1923 CTTTCCGACCCAGGAGTCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 1982
Qy 1953 AATTGAAGAGCTGTGTACAGTATCTCAGAGCTTCTCTCTATCAGTCCGCTATGCTTTC 2012
Db 1983 GACTGAAGTGTGTGACCAAGATCTCAGAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2042
Qy 2013 AGAAGCTGTGGAAGCTACCTCTTCTTTGAGACCTTGTGGAGAGGCTGTGAAATGCT 2072
Db 2043 TGAAGCTGTGGAAGCTACCTCTTCTTTGAGACCTTGTGGAGAGGCTGTGAAATGCT 2102
Qy 2073 AGGCAAGAGATAGTTTCCCGAGAGATTAAGTGTGTTGTTGGCCATCGGAGCCCAACG 2132
Db 2103 GGGCAAGAGATAGTCTCTCCAGAGATTAAGTGTGTTGTTGGCCATAGCAGCCCAACG 2162
Qy 2133 GAAGCTGAGATTCATCTCCAGAGTCCGCTCATCCAGACACCCCAATTTCCAGGAGACA 2192
Db 2163 CAAGGAGAAATTTATCCCAAGAGTCCATCTCAGACACCCAGTTCAGGAGACA 2222
Qy 2193 GAGCGGAAAGACATTTGCCAAGCCCGCTCACCATCACGCCAGCTCCATGAGGAAATCTG 2252
Db 2223 AAGTGGAGAGCTCACAGGCCCGGTCACTTACTCCAGCTCCATCAGGAGATAG 2282
Qy 2253 CACATATTTCCATAGAAAGTCCAGAGGACTTCTGTGTGCTGTGAACACTCAACAGAAAT 2312
Db 2283 TACATATCTTCTGAGAAAGACTCAAGATGACTTCTGAGTCTCTGAGTCTCTGAGAACTCACTCACTGAAT 2342
Qy 2313 ATAGATTTCTAATCTGA 2328
Db 2343 ATAGATTTCTAATCTGA 2358

RESULT 7
LOCUS DQ043568
DEFINITION Homo sapiens RAMP gene, virtual transcript, partial sequence,
genomic survey sequence.
ACCESSION DQ043568
VERSION DQ043568.1 GI:66894783

KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo
REFERENCE 1 (bases 1 to 1923)
AUTHORS Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B., Hubisz, M.J., Fledel-Alon, A., Tanenbaum, D.M., Civeello, D., White, T.J., Snieksy, J.J., Adams, M.D. and Cargill, M.
TITLE A Scan for Positively Selected Genes in the Genomes of Humans and Chimpanzees
JOURNAL (er) PLoS Biol. 3 (6), E170 (2005)
PUBMED 15869325
REFERENCE 2 (bases 1 to 1923)
AUTHORS Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B., Hubisz, M.J., Fledel-Alon, A., Tanenbaum, D.M., Civeello, D., White, T.J., Snieksy, J.J., Adams, M.D. and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment. Translation starts at the beginning of alignment.
FEATURES
source Location/Qualifiers
1..1923
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
gene <1..>1923
/gene="RAMP"
/locus_tag="HC10244"
ORIGIN
Query Match 57.5%; Score 1627.4; DB 14; Length 1923;
Best Local Similarity 86.5%; Pred. No. 0;
Matches 1812; Conservative 0; Mismatches 111; Indels 171; Gaps 1;
QY 124 ATGCTCTTCAATTCGGTGTCTCGCCAGACCCAGCTTGGGCTCTGAGAAATGGATGGTCT 183
DB 1 ATGCTCTTCAATTCGGTGTCTCGCCAGACCCAGCTTGGGCTCTGAGAAATGGATGGTCT 60
QY 184 TCACAAATACCTTCAATCCCTTCTGACTGTGATATCAGTGCAGTGGTAAATGATGAACAC 243
DB 61 TCACAAATACCTTCAATCCCTTCTGACTGTGATATCAGTGCAGTGGTAAATGATGAACAC 120
QY 244 ACTTCTTATGGAGAACAGAGTCCAGTTCCTCTCTTTGGATGTACCTTCTCTTCTGCT 303
DB 121 ACTTCTTATGGAGAACAGAGTCCAGTTCCTCTCTTTGGATGTACCTTCTCTTCTGNN 180
QY 304 CCCAATATGGAACATGACTAGCAGTTCGCAATGAAGAAGGCTTTGTCGATGTATAC 363
DB 181 NNN 240
QY 364 ACAGAAATCAAAGTTTCAGAAAGAGTCTTCAAGAAATGATGGCTCACTGGAAATGCC 423
DB 241 NNN 300
QY 424 GTCTTTGACTGGCTGGCTTCTGGTGAATTAATCTTGTACAGCAGCAGTGTATCAA 483
DB 301 GTCTTTGACTGGCTGGCTTCTGGTGAATTAATCTTGTACAGCAGCAGTGTATCAA 360
QY 484 ACAGCCAAATTTGGGACGTAAAGCTGGTGAAGCTGATTTGGAACATGCAAAAGCTCATCAA 543
DB 361 ACAGCCAAATTTGGGACGTAAAGCTGGTGAAGCTGATTTGGAACATGCAAAAGCTCATCAA 420
QY 544 TGACGCTCAAGTCAGTTCCTTTTCTAAGTTTGAGAAAGCTGTATTTCTGTCGGGTGGA 603
DB 421 TGACGCTCAAGTCAGTTCCTTTTCTAAGTTTGAGAAAG----- 460
QY 604 AGAGATGGCAACATTTATGCTGGGATACAGGTGCAACAAAGAGATGGGTTTATAGG 663

461 ----- 460
664 CAAGTGAATCAATCAGTGGAGCTCACAATACCTCAGACAAGCAACCCCTTCAAAACCC 723
461 ----- 460
724 AAGAAGAAACAGAATTCAAAGGACTTGCTCTCTTCTGTGGATTTCACGAAGTGTACT 783
461 -----GCTATCAGGATTTCCAGCAAGTGTACT 489
784 GTGGTCTCTTTCAAGACGAGAATACTTAGTCTCAGCAGAGCTGTGGATGGGATAATC 843
490 GTGGTCTCTTTCAAGACGAGAATACTTAGTCTCAGCAGAGCTGTGGATGGGATAATC 549
844 AAGATATGGGATTTACGTAAGAAATTAATCTGCTTATCGACAAGAACCCATAGCATCCAG 903
550 AAGATATGGGATTTACGTAAGAAATTAATCTGCTTATCGACAAGAACCCATAGCATCCAG 609
904 TCTTTCCGTACCCAGGTAGCAGCACTCGAAACCTTGGATATTCAGAGTCTGATTTGGAT 963
610 TCTTTCCGTACCCAGGTAGCAGCACTCGAAACCTTGGATATTCAGAGTCTGATTTGGAT 669
964 TCCACTGGCTCTACTTTATTTGCTAAATGACAGACGATACATCTACATGTTAAATG 1023
670 TCCACTGGCTCTACTTTATTTGCTAAATGACAGACGATACATCTACATGTTAAATG 729
1024 ACTGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACCAAGCACTACCTTTTAT 1083
730 ACTGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACCAAGCACTACCTTTTAT 789
1084 GTAAATCCAGCTTAGTCCAGATGACCGAGTCTTTAGTCAAGTGGTCAAGTGAAGCT 1143
790 GTAAATCCAGCTTAGTCCAGATGACCGAGTCTTTAGTCAAGTGGTCAAGTGAAGCT 849
1144 GCCTACATATGGAAGTCTCCACACCTCGCAACCTCTACTGTGCTCCTGGGTCAATCT 1203
850 GCCTACATATGGAAGTCTCCACACCTCGCAACCTCTACTGTGCTCCTGGGTCAATCT 909
1204 CAAGAGTCACTGT 1263
910 CAAGAGTCACTGT 969
1264 GATGCAATACACTAAATAATCTGGGCTTGAATAGAGCTTAGAGGAGAAACCAAGAGGT 1323
970 GATGCAATACACTAAATAATCTGGGCTTGAATAGAGCTTAGAGGAGAAACCAAGAGGT 1029
1324 GATAAATCTTCCACGCTGGGTGTGGCTCTCAGAGAAGAAAGAGTCAAGACCTGGCCTA 1383
1030 GATAAATCTTCCACGCTGGGTGTGGCTCTCAGAGAAGAAAGAGTCAAGACCTGGCCTA 1089
1384 GTAAACAGTAAACGAGTAGCAGATCTCTCTGCAAGCCCAAGGTTAAGTGAATCA 1443
1090 GTAAACAGTAAACGAGTAGCAGATCTCTCTGCAAGCCCAAGGTTAAGTGAATCA 1149
1444 TCCAAATCTTCCCGCTCATCCGAGCTTGTGCCCAAGCTGTGTGGAGACCTCCTCTCT 1503
1150 TCCAAATCTTCCCGCTCATCCGAGCTTGTGCCCAAGCTGTGTGGAGACCTCCTCTCT 1209
1504 CTTTCAAAATCTCTTACCTTCTCTTATTAATAACCTCTCTGCAAGCCCGGCTCTCCCATC 1563
1210 CTTTCAAAATCTCTTACCTTCTCTTATTAATAACCTCTCTGCAAGCCCGGCTCTCCCATC 1269
1564 AACAGAGAGGCTGTCT 1623
1270 AACAGAGAGGCTGTCT 1329
1624 ATTAGAAACTGGGTGACCCGAGCACTTCTCTCATCACCACCATCCTCTCTCTCTCTCT 1683
1330 ATTAGAAACTGGGTGACCCGAGCACTTCTCTCATCACCACCATCCTCTCTCTCTCTCT 1389
1684 GAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCATCCCAA 1743
1390 GAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCATCCCAA 1449


```
QY 1189 CTCCTGGGTCATTCTCAAGAGGTCACGCTCTGTGCTGGTGTCCATCTGACTTCACAAAG 1248
Db 781 CTCCTGGGTCATTCTCAAGAGGTCACGCTCTGTGCTGGTGTCCATCAGACTTCACCAAG 840

QY 1249 ATTGCTACCTGTTCTGTATGACAATACACTAAATACTGGCGCTTGAATAGAGGCTTAGAG 1308
Db 841 ATTGCAACCTGCTCTGATGATATACACTGAAATCTGGCGCTTGAATAGAGGCTTAGAG 900

QY 1309 GAGAACACGAGGCTGATAACTTCCACGCTGGGTTGGCGCTCTCAGAGAAAAGAG 1368
Db 901 GAGAAACC---AGGTGATAAACATTCCATAGTGGTGGACCTCTCAGAGAAAAGAA 957

QY 1369 TCAAGACCTGCGCTAGTAACAGTAACGAGTAGCCAGAGTACTCTCGCCAAAGCCCCAGG 1428
Db 958 GTGAAGCCTGCCAGGTAAACGTTACGAGTAGCCAGAGTACTCTGCAAAAGTCCAGAG 1017

QY 1429 GTAAAGTGCAATCCATCCAAATCTTCCCGTCTATCCGACGCTGTGTGCCCAAGCTGTGCT 1488
Db 1018 GCCAAGAGCAGTCCATCATCTCTCTCTCTGTCGTCAGCAGCTGTGTACTCCGAGCTGTGCA 1077

QY 1489 GGAGACCTCCTCTCTCTCAATATCTCTACCTACGTTCTCTATTAAACCTCTCTCTGCCAAG 1548
Db 1078 GGAGACCTCCTCTCTCTCAAGTACCCACACATCTCTCAGTCAAAACCACTCTCTGCCACG 1137

QY 1549 GCCCGGTCTCCCATCAACAGAGAGGCTCTGTCTCTCTCGTCTCTCCCAAGCCACCTTCA 1608
Db 1138 ACCGTTCTTCAAGTACGAGAGAGGCTCCATCTCTCTGTGCTCTCCCAAGCAGCTCTCA 1197

QY 1609 TCTTTCAAGATGTCGATTAGAACTGGGTGACCCGCAACACTTCTCTCATCACCACCCATC 1668
Db 1198 TCTTTCAAGATGTCGATTAGAACTGGGTGACCCGCAACACTTCTCTCATCACCACCTGTC 1257

QY 1669 ACTCCACCTCTCTCGGAGACCAAGATCATCTCTCCGAGAAAAGCCCTTATCTCTGTGAGC 1728
Db 1258 ACTCCACCTCTCTCGGAGACCAAGATCATCTCTCCGAGAAAAGCTTATCTCTGTGAGC 1317

QY 1729 CAGAAGTCATCCCAAGCAGAGGCTCTGTCTGAGTCTAGAAATAGAGTAAAGAGAGGCTA 1788
Db 1318 CAGAAGTCATCAGAGGAGATGCTGTCTCTGAAATCTAGAAATAGAGTAAAGAGGCTT 1377

QY 1789 GACTCAAGCTGTCTGGAGAGTGTGAACAAAGTGTGTGAAGAGTGTGTAACCTGTGTGACT 1848
Db 1378 GACTCAAGCTGTCTGGAGAGTGTGAACAAAGTGTGTGAAGAGTGTGTAACCTGTGTGACT 1437

QY 1849 GAGCTTGATGGCAAGTTGAAAATCTTTCATTGGATCTGTGCTGCTTGTGCTGTGTAACAG 1908
Db 1438 GAGCTTGACGCGCAAGCGGAGAGTCTTCTGTTGGATCTGTGCTGCTTTCGCGCACCCAG 1497

QY 1909 GAAGACCTTAGTAGGACTCTTAGGTCTTACCAGATCAAGCAAAATTTGAAGGAGGCTGTT 1968
Db 1498 GAAGTCTTAGACCAAGACTCCGAGGGTCTTACCAGATCAAGCAAAATTTGAAGGAGGCTGTT 1557

QY 1969 ACCAGTATCTCAGAGCCTCCGCTCTCTATCAGTCCGCTATGCTTTCAGAAAGCTGTGGAACG 2028
Db 1558 ACAAGATCTCAGAACCTCTCTCTGTCAGTCTCTTATGCTTCTGAGGCTGTGAGCCA 1617

QY 2029 CTACCTCTTCTCTTGAGACTTGTGGAGAAAGGCTGTGAAATGTGAGCAAAAGAGATAGT 2088
Db 1618 CTGCTCTCTCTTGAGACTTGTGGAGAAAGGATCTGAGATGTGGGCAAGAGAGATAGC 1677

QY 2089 TCCCCAGAGATAAATAACTTGTGTCATGGCCATGGCAGCCAAACGGAAGGCTGAGATCCA 2148
Db 1678 TCTCCAGAGATAAAGAACTGGTGTGTGGCCATAGCAGCCAAACGGAAGGAGAGAAATCCA 1737

QY 2149 TCTCCAGAAAGTCCGTCATCCAGACACCCAAATTCAGAGAGACAGAGCGGAAGACATTG 2208
Db 1738 TCCCAAGAAAGTCCATCATCTCAGACACCCAGTTCAGGAGACAAAGTGGGAGAGACGTCA 1797

QY 2209 CCAAGCCCGGTCCACATCAGCCCGAGCTCCATGAGGAAAATCTGCAATATCTTCCATAGA 2268
Db 1798 CCAAGCCCGGTCCACATCTCCAGCTCCATGAGGAAAGATATGTATACATCTTCTCGTAGA 1857
```

```
QY 2269 AAGTCCCGAGGAGACTTCTGTGCTCTGAACACTCAACAGAAATATAGATTTCTAAATCTGA 2328
Db 1858 AAGACTCAAGATGACTTCTGCGAGTCTGACACTCACTCACTGAATATAGATGCTAATCTGA 1917

RESULT 9
LOCUS DQ043569
DEFINITION Pan troglodytes RAMP gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION DQ043569
VERSION DQ043569.1 GI:66894784
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Pan.
REFERENCE 1 (bases 1 to 1923)
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Gnanowski,S., Sackton,T.B.,
Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.
TITLE A Scan for Positively Selected Genes in the Genomes of Humans and
Chimpanzees
JOURNAL (er) PLoS Biol. 3 (6), E170 (2005)
PUBMED 15869325
REFERENCE 2 (bases 1 to 1923)
AUTHORS Nielsen,R., Bustamante,C., Clark,A.G., Gnanowski,S., Sackton,T.B.,
Hubisz,M.J., Fledel-Alon,A., Tanenbaum,D.M., Civello,D.,
White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-2005) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment. Translation starts at the beginning of
alignment.
FEATURES
source 1..1923
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>1923
/gene="RAMP"
/locus_tag="HC10244"
ORIGIN
Query Match 51.3%; Score 1451.8; DB 14; Length 1923;
Best Local Similarity 78.3%; Pred. No. 0;
Matches 1640; Conservative 0; Mismatches 283; Indels 171; Gaps 1;
QY 124 ATGCTCTTCAATTCGGTGTCTCGCCAGCCCGAGCTTGGCGTCTGAGAAATGGATGGTCT 183
Db 1 ATGCTCTTCAATTCGGTGTCTCGCCAGCCCGAGCTTGGCGTCTGAGANNNGATGGTCT 60
QY 184 TCACAAATACCTCTTCAATCCCTTCTGACTGGTGTATCAGTGCAGTGGTATGATGAACAC 243
Db 61 TCACAAATACCTCTTCAATCGCTTCTGACTGGTGTATCAGTGCAGTGGTATGATGAACAC 120
QY 244 ACTTCTTATGGAGAAACAGGAGTCCCGAGTTCCTCTCTTTGGATGTACCTTCTTCTGCT 303
Db 121 ACTTCTTATGGAGAAACAGGAGTCCCGAGTTCCTCTCTTTGGATGTACCTTCTTCTGNN 180
QY 304 CCCAATATGGAACATGTACTAGCAGTTCGAATGAAGAGGCTTGTTCGATTGTATAAC 363
Db 181 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 240
QY 364 ACAGAAATCACAAGTTTCAGAAAGAGTCTTCAAAAGATGGATGCTCACTGGAATGCC 423
Db 241 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 300
QY 424 GTCTTTGACCTGGCCCTGGGTCTCGTGAACCTTAAACTTTGTACAGCAGCAGGTGATCAA 483
Db 301 GTCTTTGACCTGGCCCTGGGTCTCGTGAACCTTAAACTTTGTACCGCAGCAGGTGATCAA 360
```

Qy	484	ACAGCCAAATTTTGGGACGTAAAAAGCTGGTGAGCTGATTGGAAACATGCCAAAGGTCATCAA	543
Db	361	ACAGCCAAATTTTGGGACGTAAAAAGCTGGTGAGCTGATTGGAAACATGCCAAAGGTCATCAA	420
Qy	544	TGCAGCCTCAAGTCAGTGCAGTGTCTTTCTAAAGTTTCAGAAAGCTGTATTCTCTACGGGTGGA	603
Db	421	TGCAGCCTCAAGTCAGTGCAGTGTCTTTCTTTCTAAAGTTTCAGAAAG	460
Qy	604	AGAGATGGCAACATTATGGTCTGGGATACCAGGTGCAACAAAAAAGATGGTTTTATAGG	663
Db	461	-----	460
Qy	664	CAAGTGAATCAAATCAGTGGAGCTCACAATACCTCAGACAAGCAAAACCCCTTCAAAAACC	723
Db	461	-----	460
Qy	724	AAGAAGAAACAGAAATCAAAGGACTTGTCTCTTCTGTGGATTTCAGCAAGAGTGTTACT	783
Db	461	-----GCTATCAGGATTTCCAGCAAGAGTGTTACT	489
Qy	784	GTGGTCTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGAGCTGTGGATGGATAATC	843
Db	490	GTGGTCTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGAGCTGTGGATGGATAATC	549
Qy	844	AAAGTATGGGATTTACGTAAAGATTTATATCTGCTTATCGACAAGAACCCATAGCATCCAAG	903
Db	550	AAAGTATGGGATTTACGTAAAGATTTATATCTGCTTATCGACAAGAACCCATAGCATCCAAG	609
Qy	904	TCTTTCTGTACCCAGGTAGCAGCACTCGAAAACTTCGGATATTCAGTCTGTATTTGGAT	963
Db	610	TCTTTCTGTACCCAGGTAGCAGCACTCGAAAACTTCGGATATTCAGTCTGTATTTGGAT	669
Qy	964	TCCACTGGCTCTACTTTTATTTGCTAAATTGACACAGCATAACTACTACATGTTTAAATATG	1023
Db	670	TCCACTGGCTCTACTTTTATTTGCTAAATTGACACAGCATAACTACTACATGTTTAAATATG	729
Qy	1024	ACTGGTGTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACCAAGAACTCTACCTTTTAT	1083
Db	730	ACGGGTTGAAGACTTCTCCAGNNGCTATTTTCANTGGAACCAACANNCTCTACCTTTTAT	789
Qy	1084	GTAATAATCCAGCCTTAGTCCAGATGACCAGTTTTTAGTTCAGTGCCTCAAGTGATGAAGCT	1143
Db	790	GTAATAATCCAGCCTTTNNNNNATGACCAGTTTTTANNNGTGGCTCAAGTGATGAAGCT	849
Qy	1144	GCCTACATATGGAAGGTCCTCACACCCCTGGCAACCTCTACTGTGCTCCTGGGTCATTCT	1203
Db	850	GCCTACATATGGAAGGTCCTCACACCCCTGGCAACCTCTACTGTGCTCCTGGGTCATTCT	909
Qy	1204	CAAGAGTTCAGTCTGTGTGCTGTGTGTCATCTGACTTTCACAAAGATTGCTACTGTTCT	1263
Db	910	CAAGAGTTCAGTCTGTGTGCTGTGTGTCATCTGACTTTCACAAAGATTGCTACTGTTCT	969
Qy	1264	GATGCAATACACTTAAATACTCTGGCTTGAATAGAGCTTTAGAGGAGAAACACAGGAGGT	1323
Db	970	GATGCAATACACTTAAATACTCTGGCTTGAATAGAGCTTTAGAGGAGAAACACAGGAGGT	1029
Qy	1324	GATAAACTTTCCAGCGTGGGTTGGCCCTCTCAGAAGAAAAAGAGTCAAGACCTGGCCTA	1383
Db	1030	GATAAACTTTCCAGGNNGGTTGGCCCTCTCAGAAGAAAAAGAGTCAAGACCTGGCCTN	1089
Qy	1384	GTAACAGTAACGATAGCCAGAGTACTCTCTGCCAAAGCCCCAGGGTAAAGTGCAATCCA	1443
Db	1090	NN	1149
Qy	1444	TCCAAATCTTCCCCGTCAATCCGAGCTTGTGCCCAAGCTGTGCTGGAGACCTCCCTCTT	1503
Db	1150	NN	1209
Qy	1504	CTTTCAAATACTCTACGTTCTCTATTAAAACTCTCTGTCAAAGGCCCGGTCTCCCATC	1563
Db	1210	NNNNNNNNNNNNCTTACGTTCTCTATTNNAACCTCTCTGTCAAAGGCCCGGNTCCCATC	1269

QY	1564	AACAGAAAGAGGCTCTGTCTCCTCCGTCCTCTCCCAAGCCACTTTCATCTTTCAAGATGTCG	1623
Db	1270	AACAGAAAGAGGCTNTGTCTCCTCCGTCCTCTCCCAAGCCACTTTCATCTTTCAAGATNNNN	1329
QY	1624	ATTAGAACTGGGTGACCCGAAACACCTTCCTCATCACCCACCATCACTCCACCTGCTTCG	1683
Db	1330	ATTAGAACTGGGTGACCCGAAACACCTTCCTCATCACCCCATCACTCCACCTGCTTCG	1389
QY	1684	GAGACCAAGATCATGCTCCGAGAAAGCCCTATTTCCTGTGAGCCAGAAAGTCATCCCAA	1743
Db	1390	GAGACCAAGATCATGCTCCAGAAAGCCCTATTTCCTGNNAGCCAGAAAGTCATCCCAA	1449
QY	1744	GCAGAGCGTTCCTCTGAGCTAGANAATAGAGTAAGAGGAGGCTAGACTCAAGCTGCTCG	1803
Db	1450	GCAGAGCGTTCCTCTGAGCTAGANAATAGAGTAAGAGGAGGCTAGACTCAAGCTGCTCG	1509
QY	1804	GAGAGTGTGAAACAAAAAGTGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTTGATGGCCAA	1863
Db	1510	GAGAGTGTGAAACAAAAGTGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTTGATGGCCAA	1569
QY	1864	GTTGAAAAATCTTCATTTTGGATCTGTGCTGCTTCGTGTTAAACAGGAGACCTTTAGTAAG	1923
Db	1570	GTTGAAAAATCTTCATTTTGGATCTGTGCTGCTTCGTGTTAAACAGGAGACCTTTAGTAAG	1629
QY	1924	GACTCTCTAGTCCCTACCAATCAAGCAAAATTTGAAGGAGCTGTAACCACTATCTCAGAG	1983
Db	1630	GACTCTCTAGTCCCTACCAATCAAGCAAAATTTGAAGGAGCTGTAACCACTATCTCAGAG	1689
QY	1984	CCTCCGTCCTCTATCAGTCCGATGCTTCAGAAAGCTGTGGAAGCTACTCTTCCTTTG	2043
Db	1690	CCTCCGTCCTCTGTCAGTCCGATGCTTCAGAAAGCTGTGGAAGCTACTCTTCCTTTG	1749
QY	2044	AGACTTGTGGAGAAAGGCTGAAAATGGTAGGCAAGAGAAATAGTTTCCCGAGAGATAAA	2103
Db	1750	AGACTTGTGGAGAAAGGCTGAAAATGGTAGGCAAGAGAAATAGTTTCCCGAGAGATAAA	1809
QY	2104	AACGTGTTGTTGGCCATGTGGCAGCCAAACGGAAGGCTGAGAATCCATCTCCACGAAGTCGG	2163
Db	1810	AACGTGTTGTTGGCCATGTGGCAGCCAAACGGAAGGCTGAGAATCCATCTCCACGAAGTCGG	1869
QY	2164	TCATCCACAGACCCCAATTCACGAGACAGAGCGGAAAGACATTTGCCAGCCCG	2217
Db	1870	TCATCCACAGACCCCAATTCACGAGACAGAGCGGAAAGACATTTGCCAGCCCG	1923

RESULT 10
AL557830

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS
SOURCE

ORGANISM

REFERENCE

[illegible]

JOURNAL

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DJ003CC05QP1&c=592.f.

FEATURES
source
Location/Qualifiers
1..1106
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ003YF09"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 35.2%; Score 996.2; DB 1; Length 1106;
Best Local Similarity 99.4%; Pred. No. 1e-265;
Matches 1031; Conservative 0; Mismatches 3; Indels 3; Gaps 3;

QY 54 TCGATTTCGTGAACCTGGAGGCAATTTACGACTTTTCTCAGCTGAGGCTTTTC 113
DB 11 TCGGATTTCTGCTGAACCTGGAGGCAATTTCTACGACTTTTCTCAGCTGAGGCTTTTC 70
QY 114 TCCGACCTGATGCTCTTCAATTCGGTGTCCGCCAGCCAGCTTGGCGTCTCGAGAA 173
DB 71 TCCGACCTGATGCTCTTCAATTCGGTGTCCGCCAGCCAGCTTGGCGTCTCGAGAA 130
QY 174 TGGATGGTCTTCAATACCTCTTCAATCCCTTCTGACTGGTTATCAGTCAGTGGTAA 233
DB 131 TGGATGGTCTTCAATACCTCTTCAATCCCTTCTGACTGGTTATCAGTCAGTGGTAA 190
QY 234 TGAATGAACACTCTTATGGAGAAACAGAGTCCAGTTCCTCTTTGGATGTACCTT 293
DB 191 TGAATGAACACTCTTATGGAGAAACAGAGTCCAGTTCCTCTTTGGATGTACCTT 250
QY 294 CTCTTCTGCTCCCAATATGGAACATGTACTAGCAGTTCGCAATGAGAGGCTTTGTCG 353
DB 251 CTCTTCTGCTCCCAATATGGAACATGTACTAGCAGTTCGCAATGAGAGGCTTTGTCG 310
QY 354 ATTGTATAACACAGATCAAAAGTTTCAGAAAGAGTGTCTTCAAAAGATGGATGCTCA 413
DB 311 ATTGTATAACACAGATCAAAAGTTTCAGAAAGAGTGTCTTCAAAAGATGGATGCTCA 370
QY 414 CTGGAATGCGCTCTTTGACCTGGCTGGGTCTCTGGTGAACCTTAAACTTTGTTACAGCAGC 473
DB 371 CTGGAATGCGCTCTTTGACCTGGCTGGGTCTCTGGTGAACCTTAAACTTTGTTACAGCAGC 430
QY 474 AGTGATCAAAACAGCCAAATTTGGGACGTAAAGCTGTGAGCTGATTTGGAAACATGCAA 533
DB 431 AGTGATCAAAACAGCCAAATTTGGGACGTAAAGCTGTGAGCTGATTTGGAAACATGCAA 490
QY 534 AGTGATCAATGACGCTCAAGTCAGTGTGCTTTCTTAAGTTTGAGAAAGCTGATTTCTG 593
DB 491 AGTGATCAATGACGCTCAAGTCAGTGTGCTTTCTTAAGTTTGAGAAAGCTGATTTCTG 550
QY 594 TACGGGTGGAAGAGATGGCAACATTTAGTCTGGGATACAGGTGCAACAAAAAGATGG 653
DB 551 TACGGGTGGAAGAGATGGCAACATTTAGTCTGGGATACAGGTGCAACAAAAAGATGG 610
QY 654 GTTTTATAGCAAGTGAATCAATCAGTGGAGCTCAATACCTCAGACAGCAAAACCCC 713
DB 611 GTTTTATAGCAAGTGAATCAATCAGTGGAGCTCAATACCTCAGACAGCAAAACCCC 670
QY 714 TTCAAACCCACAGAGAAACAGAAATTCAAAGGACTTGTCTCTGTGGATTTCCAGCA 773
DB 671 TTCAAACCCACAGAGAAACAGAAATTCAAAGGACTTGTCTCTGTGGATTTCCAGCA 730
QY 774 AAGTGTTACTGTGCTCTCTTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGA 833
DB 731 AAGTGTTACTGTGCTCTCTTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGA 790

QY 834 TGGGATAATCAAAAGTATGGGATTTAGCTAAGAAATTATATCTGCTTATCGACAAGAACCCAT 893
DB 791 TGGGATAATCAAAAGTATGGGATTTAGCTAAGAAATTATATCTGCTTATCGACAAGAACCCAT 850
QY 894 AGCATCCAAGTCTTTCTCTGTACCCAGGTAGCAGCACTCGAAAACCTT-GGATATTCAAGTC 952
DB 851 AGCATCCAAGTCTTTCTCTGTACCCAGGTAGCAGCACTCGAAAACCTTGGGATATTCAAGTC 910
QY 953 TGATTTTGAAATCCCACTGGCTCTACTTT-ATTGCTAATTGCACAGACATAACATCTAC 1011
DB 911 TGATTTTGAAATCCCACTGGCTCTACTTTTAAATTTGCTAATTGCACAGACATAACATCTAC 970
QY 1012 ATGTTTAAATATGACTGGGTGAAGACTTCTCAGTGGCTATTTTCAATGGACACCAAGAAC 1071
DB 971 ATGTTTAAATATGACTGGGTGAAGACTTCTCAGTGGCTATTTTCAATGGACACCAAGAAC 1029
QY 1072 TCTACCTTTTATGTAAA 1088
DB 1030 TCTACCTTTTATGTAAA 1046

RESULT 11
AL553783
LOCUS
DEFINITION AL553783 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0DI077YK11 5-PRIME, mRNA sequence.
ACCESSION AL553783
VERSION AL553783.3 GI:45858548
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 998)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:31275597.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Crémieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and EcoR V
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1..998
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI077YK11"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 35.1%; Score 992.8; DB 1; Length 998;
Best Local Similarity 99.0%; Pred. No. 8.7e-265;
Matches 988; Conservative 9; Mismatches 1; Indels 0; Gaps 0;

QY 1527 TATTAACCTCTCTGCCAAGCCGGTCTCCCATCAACAGAGAGGCTCTGTCTCTC 1586
DB 1 TATTAACCTCTCTGCCAAGCCGGTCTCCCATCAACAGAGAGGCTCTGTCTCTC 60
QY 1587 CGTCTCTCCCAAGCACCTTTCATCTTCAAGATGTGATTAGAACTGGGTGACCCGAAC 1646

```
Db 61 CGTCTCTCCCAAGCCACCTTCATCTTCAAGATGTCGATTAGAAAGTGGTGACCCGAAAC 120
QY 1647 ACCTTCCTCATCACCACCCATCCTCCACCTCGTTCGGAGACCAAGATCATGTCTCCGAG 1706
Db 121 ACCTTCCTCATCACCACCCATCCTCCACCTCGTTCGGAGACCAAGATCATGTCTCCGAG 180
QY 1707 AAAAGCCCTTATTCCTGTGAGCGAGAGTCATCCCAAGCAGAGGCTTGCTCTGAGTCTAG 1766
Db 181 AAAAGCCCTTATTCCTGTGAGCGAGAGTCATCCCAAGCAGAGGCTTGCTCTGAGTCTAG 240
QY 1767 AAATAGAGTAAAGAGGAGCTAGACTCAAGCTGTCTGGAGAGTGTCAAAACAAAAGTGTGT 1826
Db 241 AAATAGAGTAAAGAGGAGCTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAAGTGTGT 300
QY 1827 GAAGAGTTGTAACTGTGTGACTGAGCTTGATGGCCAAAGTTGAAATCTTTCATTTGGATCT 1886
Db 301 GAAGAGTTGTAACTGTGTGACTGAGCTTGATGGCCAAAGTTGAAATCTTTCATTTGGATCT 360
QY 1887 GTGCTGCCTTGTGTTAACCCAGAGACCTTAGTAAGGACTCTTAGGTCTTACCAATC 1946
Db 361 GTGCTGCCTTGTGTTAACCCAGAGACCTTAGTAAGGACTCTTAGGTCTTACCAATC 420
QY 1947 AAGCAAAATTGAAGGAGCTGTGTACCAAGTATCTCAGAGCCTCCGTCTCTCTATCAGTCCGTA 2006
Db 421 AAGCAAAATTGAAGGAGCTGTGTACCAAGTATCTCAGAGCCTCCGTCTCTCTATCAGTCCGTA 480
QY 2007 TGCCTCAGAAAGCTGTGGAAAGCTACCTCTTCTTTGAGACCTTGTGGAGAAAGGTCTGA 2066
Db 481 TGCCTCAGAAAGCTGTGGAAAGCTACCTCTTCTTTGAGACCTTGTGGAGAAAGGTCTGA 540
QY 2067 AATGTTAGGCAAGAGAAATAGTTCCCCAGAGATTAAGGACTCTTAGGTCTTGGCCATGCGCAGC 2126
Db 541 AATGTTAGGCAAGAGAAATAGTTCCCCAGAGATTAAGGACTCTTAGGTCTTGGCCATGCGCAGC 600
QY 2127 CAAAAGGAGGCTGAGAAATCCATCTCCAGAAAGTCCGTATCCAGACACCCAAATCCAG 2186
Db 601 CAAAAGGAGGCTGAGAAATCCATCTCCAGAAAGTCCGTATCCAGACACCCAAATCCAG 660
QY 2187 GAGACAGCGGAAAGACATTCGCAAGCCCGTCAACCATCAGCCGAGCTCCATGAGGAA 2246
Db 661 GAGACAGCGGAAAGACATTCGCAAGCCCGTCAACCATCAGCCGAGCTCCATGAGGAA 720
QY 2247 AATCTGCACATCTTCCATAGAAAGTCCAGAGGACTTCTGTGCTCTGACACTCAAC 2306
Db 721 AATCTGCACATCTTCCATAGAAAGTCCAGAGGACTTCTGTGCTCTGACACTCAAC 780
QY 2307 AGAATTTATAGATTTCTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTMAAAACAAGCTGAG 2366
Db 781 AGAATTTATAGATTTCTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTMAAAACAAGCTGAG 840
QY 2367 CTTTGGTCCACTAAACCAAGATGAAATAACAAGAGTGACTCTATAAATCTCTGGTCTTTAA 2426
Db 841 CTTTGGTCCACTAAACCAAGATGAAATAACAAGAGTGACTCTATAAATCTCTGGTCTTTAA 900
QY 2427 GAAAGCTGCTTTCTAATTTTGAAGAAATCTTTCAAGCGCTGAAATGTACCTAATCTGG 2486
Db 901 GAAAGCTGCTTTCTAATTTTGAAGAAATCTTTTCAAGCGCTGAAATGTACCTAATCTGG 960
QY 2487 TTCTACTACATAATATATATGAGCTTCCCGAGGATG 2524
Db 961 TTCTACTACATAATATGATGAGCTTCCCGAGGATG 998
```

```
RESULT 12
AL576790/c 1042 bp mRNA linear EST 07-APR-2004
LOCUS
DEFINITION AL576790 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
Clone CS0D107YK11 3-PRIME, mRNA sequence.
ACCESSION AL576790
VERSION AL576790.3 GI:46255889
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchntoglires; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1042)
Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 16, 2001 this sequence version replaced gi:31315071.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dt) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES

Location/Qualifiers

source

1..1042

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clones="CS0D107YK11"

/tissue type="PLACENTA COT 25-NORMALIZED"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dt)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

```
Query Match 34.5%; Score 976.4; DB 1; Length 1042;
Best Local Similarity 97.7%; Pred. No. 3.3e-260; Indels 2; Gaps 2;
Matches 1000; Conservative 9; Mismatches 13;

QY 1767 AAATAGAGTAAAGAGGAGCTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAAGTGTGT 1826
Db 1022 AGAATRGRTTAAGRGCGCTAACTCAAGCTTCTGGAGAGTG-KAAACAAAAGTGTGT 964
QY 1827 GAAGAGTTGTAACTGTGTGACTGAGCTTGTATGGCCAAAGTTGAAATCTTTCATTTGGATCT 1886
Db 963 GAAGAG-TGTTAACTGTGTGACTGAGCTTGTATGSCCAAGTTGAAATCTTTCATTTGGATCT 905
QY 1887 GTGCTGCCTTCTGTTAACCCAGGAGACCTTAGTAAGGACTCTTAGGTCTTACCAATC 1946
Db 904 GTGCTGCCTTCTGTTAACCCAGGAGACCTTAGTAAGGACTCTTAGGTCTTACCAATC 845
QY 1947 AAGCAAAATTGAAGGAGCTGGTACCAGTATCTCAGAGGCTCCGTCTCTCTATCAGTCCGTA 2006
Db 844 AAGCAAAATTGAAGGAGCTGGTACCAGTATCTCAGAGGCTCCGTCTCTCTATCAGTCCGTA 785
QY 2007 TGCCTCAGAAAGCTGTGAAACGCTACCTCTTCTTTGAGACCTTGTGGAGAAAGGTCTGA 2066
Db 784 TGCCTCAGAAAGCTGTGAAACGCTACCTCTTCTTTGAGACCTTGTGGAGAAAGGTCTGA 725
QY 2067 AATGTTAGGCAAGAGAAATAGTTCCCCAGAGATTAAGGACTCTTAGGTCTTGGCCATGCGCAGC 2126
Db 724 AATGTTAGGCAAGAGAAATAGTTCCCCAGAGATTAAGGACTCTTAGGTCTTGGCCATGCGCAGC 665
QY 2127 CAAAAGGAGGCTGAGAAATCCATCTCCACGAAGTCCGTATCCAGACACCCAAATCCAG 2186
Db 664 CAAAAGGAGGCTGAGAAATCCATCTCCACGAAGTCCGTATCCAGACACCCAAATCCAG 605
QY 2187 GAGACAGCGGAAAGACATTCGCAAGCCCGTACCATCAGCCGAGCTCCATGAGGAA 2246
Db 604 GAGACAGCGGAAAGACATTCGCAAGCCCGTACCATCAGCCGAGCTCCATGAGGAA 545
QY 2247 AATCTGCACATCTTCCATAGAAAGTCCAGAGGACTTCTGTGCTCTGACACTCAAC 2306
Db 544 AATCTGCACATCTTCCATAGAAAGTCCAGAGGACTTCTGTGCTCTGACACTCAAC 485
QY 2307 AGAATTTATAGATTTCTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTMAAAACAAGCTGAG 2366
```

Db 484 AGAATTATAGATTCTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTAAAAACAAGCTGAG 425
QY 2367 CTTTGTGTCACCTAAACAGATGAAATAACAGAGTGAGTCTATTAACCTCTGGTCTTTAA 2426
Db 424 CTTTGTGTCACCTAAACAGATGAAATAACAGAGTGAGTCTATTAACCTCTGGTCTTTAA 365
QY 2427 GAAAGCTGCTCTTTTCAATTTTATAGACAAAATCTTTTCAAGCTGGAATGTACCTAAATCTGG 2486
Db 364 GAAAGCTGCTCTTTTCAATTTTATAGACAAAATCTTTTCAAGCTGGAATGTACCTAAATCTGG 305
QY 2487 TTCTACTACATAATGATATGATGAGCTTCCGAGGATGAATGCTGTGTTAAATTTTATA 2546
Db 304 TTCTACTACATAATGATATGATGAGCTTCCGAGGATGAATGCTGTGTTAAATTTTATA 245
QY 2547 AAGTAATTTGTCACTCTAGCATTTTGAATGATAGTCTTCACTTTTAAATTTTATTCATC 2606
Db 244 AAGTAATTTGTCACTCTAGCATTTTGAATGATAGTCTTCACTTTTAAATTTTATTCATC 185
QY 2607 TTCTCTATATATGATGATGATGAGCTTCCGAGGATGAATGCTGTGTTAAATTTTATA 2666
Db 184 TTCTCTATATATGATGATGATGAGCTTCCGAGGATGAATGCTGTGTTAAATTTTATA 125
QY 2667 AAATTTCTATGCTAGTGAAGTATCTGCCAGCCACAGATGAGGCTGTGAAGGCTG 2726
Db 124 AAATTTCTATGCTAGTGAAGTATCTGCCAGCCACAGATGAGGCTGTGAAGGCTG 65
QY 2727 ACTGAGAAATCTCTGCTGAGACCCCTGGTCTGTTCTGCTCCCAACATGATATATTT 2786
Db 64 ACTGAGAAATCTCTGCTGAGACCCCTGGTCTGTTCTGCTCCCAACATGATATATNKD 5
QY 2787 ATTT 2790
Db 4 NTTT 1

RESULT 13
BX416133
LOCUS
DEFINITION BX416133 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
CS0DA002Y114 5-PRIME, mRNA sequence.
ACCESSION BX416133
VERSION BX416133.2 GI:46926733
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 1075)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 13, 2003 this sequence version replaced gi:30641768.
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1..1075
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DA002Y114"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into

the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

Query Match 32.8%; Score 927.2; DB 4; Length 1075;
Best Local Similarity 99.6%; Pred. No. 1.8e-246;
Matches 948; Conservative 2; Mismatches 0; Indels 2; Gaps 2;

ORIGIN

QY 1043 CAGTGGCTATTTCATATGAGACACACAGAACTCTACCTTTTATGTAAATCCACGCTTAGTC 1102
Db 1 CAGTGGCTATTTCATATGAGACACACAGAACTCTACCTTTTATGTAAATCCACGCTTAGTC 60
QY 1103 CAGATGACCAAGTTTATGTCAGTGGCTCAAGTGATGAAGCTCCCTACATATGAAAGTCT 1162
Db 61 CAGATGACCAAGTTTATGTCAGTGGCTCAAGTGATGAAGCTCCCTACATATGAAAGTCT 120
QY 1163 CCACACCTGGCAACCTCTACTGTCTCTGGGTCAATCTCAAGAGTCAAGTCTGTGT 1222
Db 121 CCACACCTGGCAACCTCTACTGTCTCTGGGTCAATCTCAAGAGTCAAGTCTGTGT 180
QY 1223 GCTGTGTCCATCTGACTTTCACAAAGATTGCTACCTGTTCTGATGACAAATACATAAAAA 1282
Db 181 GCTGTGTCCATCTGACTTTCACAAAGATTGCTACCTGTTCTGATGACAAATACATAAAAA 240
QY 1283 TCTGGCGCTTGAATAGAGGCTTTAGAGGAGAAACACAGGAGGTGATAACTTTCCACGGTGG 1342
Db 241 TCTGGCGCTTGAATAGAGGCTTTAGAGGAGAAACACAGGAGGTGATAACTTTCCACGGTGG 300
QY 1343 GTTGGCGCTCTCAGAAAGAAAAAGAGTCAAGACCTTGGCTTAGTAAACAGTAACGAGTAGCC 1402
Db 301 GTTGGCGCTCTCAGAAAGAAAAAGAGTCAAGACCTTGGCTTAGTAAACAGTAACGAGTAGCC 360
QY 1403 AGAGTACTCTGCGCAAGCCCGCCAGGGTAAAGTGCAATCCATCAATCTTCCCGTCAAT 1462
Db 361 AGAGTACTCTGCGCAAGCCCGCCAGGGTAAAGTGCAATCCATCAATCTTCCCGTCAAT 420
QY 1463 CGCGAGCTTGTGCCCAAGCTGTGTGGAGAGCTCCCTCTCTCTTCAATACTCTCTACGT 1522
Db 421 CGCGAGCTTGTGCCCAAGCTGTGTGGAGAGCTCCCTCTCTTCAATACTCTCTACGT 480
QY 1523 TCTCTATTAATAACCTCTCTGCAAGGCGGTCTCCCATCAACAGAGAGGCTCTGTCT 1582
Db 481 TCTCTATTAATAACCTCTCTGCAAGGCGGTCTCCCATCAACAGAGAGGCTCTGTCT 540
QY 1583 CTTCCGTCTCTCCCAAGCCACCTTCATCTTTCAAGATGTCGATTAGAACTCGGTGACCC 1642
Db 541 CTTCCGTCTCTCCCAAGCCACCTTCATCTTTCAAGATGTCGATTAGAACTCGGTGACCC 600
QY 1643 GAAACACCTTCTCATCACCCCATCATCCACCTGCTTCGGAGACCAAGATCATGTCTC 1702
Db 601 GAAACACCTTCTCATCACCCCATCATCCACCTGCTTCGGAGACCAAGATCATGTCTC 660
QY 1703 CGAGAAAAGCCCTTATTCCTGTGAGCCAGAAAGTCAATCCCAAGCAGAGGCTTGTCTGAGT 1762
Db 661 CGAGAAAAGCCCTTATTCCTGTGAGCCAGAAAGTCAATCCCAAGCAGAGGCTTGTCTGAGT 720
QY 1763 CTAGAAAATAGATAAGAGGAGGCTTAGACTCAAGCTGTCTGAGAGTGTAACAAAGCT 1822
Db 721 CTAGAAAATAGATAAGAGGAGGCTTAGACTCAAGCTGTCTGAGAGTGTAACAAAGCT 779
QY 1823 GTGTGAAGAGTTGTAACCTGTGTGATGTAGCTTGTAGTGCCCAAGTTGAAAATCTTCATTGG 1882
Db 780 GTGTGAAGAGTTGTAACCTGTGTGATGTAGCTTGTAGTGCCCAAGTTGAAAATCTTCATTGG 839
QY 1883 ATCTGTGCTGCTGCTGGTAAACAGGAGACCTTAGTAAGAGCTCTCTAGGTCTTACCA 1942
Db 840 ATCTGTGCTGCTGCTGGTAAACAGGAGACCTTAGTAAGAGCTCTCTAGGTCTTACCA 899
QY 1943 AATCAAGCAAAATTAAGAGAGCTGTGTACAGTATCTCAGAGGCTCGCTCTCC 1994
Db 900 AATCAAGCAAAATTAAGAGAGCTGTGTAGTATCTCAGAGGCTCGCTCTCY 950

CONSRMT
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS

RIKEN Genome Exploration Research Group
Antisense transcription in the mammalian transcriptome
Science 309 (5740), 1564-1566 (2005)
16141073

8 (bases 1 to 4308)
Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K.,
Hori, F., Iida, J., Imamura, K., Imotani, K., Itoh, M., Kanagawa, S.,
Kawai, J., Kojima, M., Konno, H., Murata, M., Nakamura, M., Ninomiya, N.,
Nishiyori, H., Nomura, K., Ohno, M., Sakazume, N., Sano, H., Sasaki, D.,
Shibata, K., Shiraki, T., Tagami, W., Tagami, Y., Waki, K., Watahiki, A.,
Muramatsu, M., and Hayashizaki, Y.

Direct Submission
Submitted (14-APR-2004) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration and Chemical Research, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa, 230-0045, Japan (E-mail: genome-res@gsr.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers
1. 4308
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:E130114J22"
/db_xref="taxon:10090"
/clone="E130114J22"
/tissue_type="eyeball"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
misc_feature
1. 4308
/notes="L2DTL protein homolog [Mus musculus]
(UniProt|Q8BW38, evidence: FASTV, 91.9%ID, 52.1%length,
match=1145)
putative"

ORIGIN

Query Match 31.4%; Score 890; DB 6; Length 4308;
Best Local Similarity 87.1%; Pred. No. 6.1e-236;
Matches 1002; Conservative 0; Mismatches 140; Indels 8; Gaps 2;

QY 12 GAGTTGGAGCGGTAACGATTGTTGTGTGAGAGCGCAACGTGCGATTCTCTGCTGAAT 71
DB 8 GAGTTGGAGCGGTAACGATTGTTGTGTGAGAGCGGCGGTTCTCTGCTGAAT 67

QY 72 TGGAGGC-ATTCTAGACATTTTCTCTCAGCTGAGGCTTTCTCTCGACCGCTGATGCTCT 130
DB 68 TGGAGCGGATCTTCTCGGCGTTCCCTCAGCTGAGGCTTCTCTCAGGCGGATGCTCT 127

QY 131 TCAATTGGGTGTCGCGCAGCCAGCTTGGCGTCTCTGAGAAATGAGTGTTCACAT 190
DB 128 TCAACTCGGTACTCGCGCAGCCAGCTGCGCGTCTCTGAGAAACGGGTGCTTTCACAT 187

QY 191 ACCCTCTCAATCCCTCTGACTGGTTATCAGTGCAGTGGTATGATGACACATCTT 250
DB 188 ACCCTCTCAATCCCTCTGAGTGGTTATCAGTGCAGTGGTATGATGACACATCTT 247

QY 251 ATGAGAAACAGGAGTCCCAAGTTCCTCTTTGGATGATACCTTCTCTTCTGCTCCCAATA 310
DB 248 ATGAGAAACAGGAGTCCCAAGTTCCTCTTTGGATGATACCTTCTGACGCTCCAGTA 307

QY 311 TGAACATGATPACGAGTTGCCAATGAGAGCGCTTTGTCGATTGTATTAACACAGAT 370
DB 308 TGGAGCATATATTAGCAGTTGCTAATGAGAGCGCTTTGTCGATTGTATTAACACAGAT 367

QY 371 CACAAAGTTTCAGAAAGAGTGTCTTCAAGAAATGGATGGCTCACTGGAATGCGTCTTTG 430
DB 368 CACAAACTAGCAAAAGAGCATGCTTCAAGAGTGGTGGCTCACTGGAATGCTGCTTTG 427

QY 431 ACTGCGCTGGGTTCCTGGTGAACTTTAACTTTTACAGCAGCAGGTGATCAAAACAGCCA 490
DB 428 ACTTGGCTGGGTTCCTGGTGAACTTTAACTTTTACAGCAGCAGGTGATCAAAACAGCCA 480

QY 491 AATTTTGGGACCTTAAAGCTGCTGAGCTGATTTGGAAACATGCAAGAGTCAATCAATGCAGCC 550
DB 481 AATTTTGGGATGTAAAGCTGCTGAGCTGATTTGGAAACATGCAAGAGTCAATCAATGCAGCC 540

QY 551 TCAAGTCAGTTCCTTTCTAAGTTTGAAGAGCTGTATTTCTGACGGTGGAAAGATG 610
DB 541 TCAAGTCTGTAGCTTTCCCAAGTTTCAAAAGCTGTGTTCTTACAGGGGGAGAGAGC 600

QY 611 GCAACATTATGTCCTGGGATACAGGTGCAACAAAAGATGGTTTTATAGCAAGTGA 670
DB 601 GCAACATTATGTCCTGGGATACAGGTGCAACAAAAGATGGTTTTATAGCAAGTGA 660

QY 671 ATCAATCAGTCGAGCTCACAATACCTCAGACAAGCAACCCCTTCARAAACCAAGAGA 730
DB 661 ATCAATCAGTCGAGCTCACAATACCTCAGACAAGCAACCCCTTCARAAACCAAGAGA 720

QY 731 AACAGAAATCAAAAGGACTTGTCTCTCTGTGATTTCCAGCAAAAGTGTACTGTGCTCC 790
DB 721 AACAAATTCAAAAGGACTTGTCTCTCTGTGATTTCCAGCAGAGTGTACTGTGCTCC 780

QY 791 TCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATTAATCAAGATAT 850
DB 781 TCTTTCAAGATGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATTAATCAAGATAT 840

QY 851 GGGATTTCAGTAAGAAATATACCTGCTTTCGACAAGAACCCCTAGCATCCAAAGTCTTTCC 910
DB 841 GGGATTTCGCAAGAAATATACCTGCTTTCGACAAGAACCCCTAGCATCCAAAGTCTTTCC 900

QY 911 TGTACCCAGTACGAGCACTCGAAAACCTTGGATATTTCAAGTCTGATTTTGGATTTCCACTG 970
DB 901 TGTACCCAGTACGAGCACTCGAAAACCTTGGATATTTCAAGTCTGATTTTGGATTTCCACTG 960

QY 971 GCTCTACTTTATTTGCTAATTCGACAGCATTAACATCTACATGTTTATATGACTGGGT 1030
DB 961 GCTCTACTTTATTTGCTAATTCGACAGCATTAACATCTATATGTTTCAATATGACTGGGT 1020

QY 1031 TGAAGACTTCTCCAGTGGCTATTTTCAATGGACCCAGCACTTACCTTTTATGTAAT 1090
DB 1021 TAAAGACTTCTCCGTTGGCTGTCTTCAATGGACCCAGCACTTACCTTTTATGTAAT 1080

QY 1091 CCAGCTTAGTCCAGATGACCACTTTTATGTCAGTGGCTCAAGTGTGATGAAGCTGCCTACA 1150
DB 1081 CAAGCTTAGTCCAGATGACCACTTTTATGTCAGTGGTTCAGTGTGATGAAGCTGCCTACA 1140

QY 1151 TATGGAAGT 1160
DB 1141 TTTGGAAGT 1150

RESULT 15
BQ431493
LOCUS
DEFINITION
BQ431493
5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE

BQ431493
AGENCOURT 7831396 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6152134
5', mRNA sequence.
BQ431493
BQ431493.1 GI:21170579
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 949)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Prepared by: Life Technologies, Inc.

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM13489 row: h column: 23

High quality sequence stop: 639.

FEATURES
source

1. .949
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6152134"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 67"
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."

ORIGIN

Query Match 31.3%; Score 885.8; DB 3; Length 949;
Best Local Similarity 98.2%; Pred. No. 6e-235;
Matches 927; Conservative 0; Mismatches 13; Indels 4; Gaps 3;

QY 1465 GCAGCTTGCGCCCAAGCTGTGTGAGACCTCCCTCTTCTTCAAAATCTCTAGGTTTC 1524
DB 1 GCAGCTTGCGCCCAAGCTGTGTGAGACCTCCCTCTTCTTCAAAATCTCTAGGTTTC 60

QY 1525 TCTATTAAACCTCTCTCCAGGCGGTCTCCCATCAAGAGAGGCTCTGTCTCC 1584
DB 61 TCTATTAAACCTCTCTCCAGGCGGTCTCCCATCAAGAGAGGCTCTGTCTCC 120

QY 1585 TCCGTCTCTCCCAAGCAGCTTTCATCTTTCAAGATGTCGATTAGAACTGGGTGACCCGA 1644
DB 121 TCCGTCTCTCCCAAGCAGCTTTCATCTTTCAAGATGTCGATTAGAACTGGGTGACCCGA 180

QY 1645 ACACCTTCTCTATCACACCCATCATCTCCACTGCTTCGGAGACCAAGATCATGTCTCG 1704
DB 181 ACACCTTCTCTATCACACCCATCATCTCCACTGCTTCGGAGACCAAGATCATGTCTCG 240

QY 1705 AGAAAGCCCTTATCTCTGTGACCAAGAGTTCATCCCAAGCAGAGGCTGTCTGAGTCT 1764
DB 241 AGAAAGCCCTTATCTCTGTGACCAAGAGTTCATCCCAAGCAGAGGCTGTCTGAGTCT 300

QY 1765 AGAAATAGAGTAAGAGGAGGTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAGTGT 1824
DB 301 AGAAATAGAGTAAGAGGAGGTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAGTGT 360

QY 1825 GTGAAGAGTTGTAACCTGTGTGACTGAGCTTGTATGGCCAAAGTTGAAATCTTATTGGAT 1884
DB 361 GTGAAGAGTTGTAACCTGTGTGACTGAGCTTGTATGGCCAAAGTTGAAATCTTATTGGAT 420

QY 1885 CTGTGCTGCTTGTCTGTTAAACAGGAGACCTTGTAGTAAAGTCTCTAGGCTCTTACCAAA 1944
DB 421 CTGTGCTGCTTGTCTGTTAAACAGGAGACCTTGTAGTAAAGTCTCTAGGCTCTTACCAAA 480

QY 1945 TCAAGCAAAATTTGAAGAGAGTGGTACAGTATCTCAGAGCCTCCGCTCTCTATCAGTCCG 2004
DB 481 TCAAGCAAAATTTGAAGAGAGTGGTACAGTATCTCAGAGCCTCCGCTCTCTATCAGTCCG 540

QY 2005 TATGCTTTCAGAAAGCTGTGGAACGCTTACCTTCTTCTTTGAGACCTTGTGGAGAGGTTCT 2064
DB 541 TATGCTTTCAGAAAGCTGTGGAACGCTTACCTTCTTCTTTGAGACCTTGTGGAGAGGTTCT 600

QY 2065 GAAATCGTAGGCAAGAGAAATAGTCTCCCAAGAGAAATAAAGTGTGTTGGCCATGGCA 2124

DB 601 GAAATGGTAGGCAAGAGAAATAGTTCCCGAGAGAAATAAAGTGTGTTGGCCATGGCA 660
QY 2125 GCCAAACGGAAGGCTGAGAAATCCATCTCCAGAAAGTCCGTATCCAGACACCCAAATTC 2184
DB 661 GCCAAACGGAAGGCTGAGAAATCCATCTCCAGAAAGTCCGTATCCAGACACCCAAATTC 720
QY 2185 AGGAGACAGAGCGGGAAGACATTCGCAAGCCCGGTCAACATCAGCCCGAGCTCCATGAGG 2244
DB 721 AGGAGACAGAGCGGGAAGACATTCGCAAGCCCGGTCAACATCAGCCCGAGCTTCATGANG 780
QY 2245 AAAATCTGCACATCTCTCCATAGAAAGT-CCAGAGGAGTCTTCTGTGTTCTCTGAACACTC 2303
DB 781 AAAATCTGCACATCTCTCATAGAAAGTCCCGAGGAGACTTCTGTGTTCTCTGAACACTC 840
QY 2304 AACAGAAATATAGATTTCTAATCTCTGAGTGAGTTACT-GAGCTTTGTTGTTCAAAACAAGC 2362
DB 841 AACAGAAATATAGATTTCTAATCTCTGAGTGAGTTACTCTGGAGCTTTGTTGTTCAAAACAAGC 900

QY 2363 TGA--GCTTTGTTGTTCAAAACAAGATGAAATAATACAAGAGTG 2404
DB 901 TGAAGCTTTGGTCCACTTAAACAGATGGAATAATACCAGAGTG 944

Search completed: November 9, 2006, 09:59:09
Job time : 13056 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2006, 06:20:51 ; Search time 643 Seconds

(without alignments)
8714.811 Million cell updates/sec

Title: US-10-726-160-1

Perfect score: 2831

Sequence: 1 ggcacagcgggagttggag.....atgaaaaaaaaaaaaa 2831

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2471159 seqs, 989689746 residues

Total number of hits satisfying chosen parameters: 4942318

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA New.*

- 1: /EMC Celerra_SIDS3/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 2: /EMC Celerra_SIDS3/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 3: /EMC Celerra_SIDS3/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 4: /EMC Celerra_SIDS3/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 5: /EMC Celerra_SIDS3/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 6: /EMC Celerra_SIDS3/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 7: /EMC Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 8: /EMC Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 9: /EMC Celerra_SIDS3/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 10: /EMC Celerra_SIDS3/ptodata/2/pubpna/US60_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2831	100.0	2831	8	US-11-266-748A-23628
2	2789.8	98.5	4221	8	US-11-266-748A-23288
3	2789.8	98.5	4221	8	US-11-266-748A-56960
4	2442.4	86.3	4083	6	US-10-540-898-673
5	1937.4	68.4	1949	8	US-11-266-748A-184890
6	1302.2	46.0	2436	6	US-10-540-898-670
7	1235	43.6	1270	8	US-11-266-748A-184891
8	1235	43.6	1270	8	US-11-266-748A-192385
9	835.8	29.5	8892	6	US-10-540-898-672
10	802	28.3	931	8	US-11-266-748A-93508
11	802	28.3	931	8	US-11-266-748A-146319
12	799.4	28.2	883	8	US-11-266-748A-41156
13	799.4	28.2	883	8	US-11-266-748A-210423
14	700.8	24.8	798	8	US-11-266-748A-261935
15	700.8	24.8	798	8	US-11-266-748A-322452
16	649.8	23.0	57860	6	US-10-540-898-669
17	613	21.7	745	6	US-10-540-898-675
18	495.2	17.5	500	8	US-11-266-748A-9970
19	439.4	15.5	485	8	US-11-266-748A-262502
20	439.4	15.5	485	8	US-11-266-748A-323019
21	291.4	10.3	464	6	US-10-513-369-1049
22	275.8	9.7	379	8	US-11-266-748A-232873
23	103.6	3.7	1914	9	US-11-218-305-21058

24	54.4	1.9	767	6	US-10-449-902-13367	Sequence 13367, A
25	48.2	1.7	1548	8	US-11-217-529-75767	Sequence 75767, A
26	43.2	1.5	8404	6	US-10-517-441-258	Sequence 258, App
27	43.2	1.5	8404	6	US-10-517-441-532	Sequence 532, App
28	43.2	1.5	8404	6	US-10-506-111-144	Sequence 144, App
29	43.2	1.5	8404	6	US-10-506-111-272	Sequence 272, App
30	40.6	1.4	1000	8	US-11-266-748A-202830	Sequence 202830, A
31	40.2	1.4	831	8	US-11-266-748A-178827	Sequence 178827, A
32	40.2	1.4	988	8	US-11-266-748A-85525	Sequence 85525, A
33	40.2	1.4	988	8	US-11-266-748A-138336	Sequence 138336, A
34	40.2	1.4	1362	8	US-11-266-748A-228469	Sequence 228469, A
35	40.2	1.4	1522	8	US-11-266-748A-25251	Sequence 25251, A
36	40.2	1.4	1582	8	US-11-266-748A-257849	Sequence 257849, A
37	40.2	1.4	1582	8	US-11-266-748A-318366	Sequence 318366, A
38	40.2	1.4	1785	8	US-11-266-748A-355179	Sequence 355179, A
39	40.2	1.4	1785	8	US-11-266-748A-385180	Sequence 385180, A
40	40.2	1.4	1785	8	US-11-266-748A-438558	Sequence 438558, A
41	40.2	1.4	2154	6	US-10-449-902-21862	Sequence 21862, A
42	40	1.4	5344	8	US-11-266-748A-31052	Sequence 31052, A
43	39.8	1.4	2625	9	US-11-056-355B-78160	Sequence 78160, A
44	39.8	1.4	2769	9	US-11-056-355B-75238	Sequence 75238, A
45	39.8	1.4	3350	9	US-11-056-355B-108274	Sequence 108274, A

ALIGNMENTS

RESULT 1

US-11-266-748A-23628

; Sequence 23628, Application US/11266748A

; Publication No. US20060134663A1

; GENERAL INFORMATION:

; APPLICANT: Harkin, Paul

; APPLICANT: Johnston, Patrick

; APPLICANT: Mulligan, Karl

; TITLE OF INVENTION: Transcriptome Microarray Technology and

; FILE REFERENCE: 55815-0102 (319189)

; CURRENT APPLICATION NUMBER: US/11/266, 748A

; PRIOR FILING DATE: 2005-11-03

; PRIOR APPLICATION NUMBER: EP 04105479.2

; PRIOR FILING DATE: 2004-11-03

; PRIOR APPLICATION NUMBER: EP 04105482.6

; PRIOR FILING DATE: 2004-11-03

; PRIOR APPLICATION NUMBER: EP 04105483.4

; PRIOR FILING DATE: 2004-11-03

; PRIOR APPLICATION NUMBER: EP 04105507.0

; PRIOR FILING DATE: 2004-11-03

; PRIOR APPLICATION NUMBER: EP 04105485.9

; PRIOR FILING DATE: 2004-11-03

; PRIOR APPLICATION NUMBER: EP 04105484.2

; PRIOR FILING DATE: 2004-11-03

; PRIOR APPLICATION NUMBER: US 60/662,276

; PRIOR FILING DATE: 2005-03-14

; PRIOR APPLICATION NUMBER: US 60/700,293

; PRIOR FILING DATE: 2005-07-18

; NUMBER OF SEQ ID NOS: 483996

; SOFTWARE: Patent in version 3.3

; SEQ ID NO 23628

; LENGTH: 2831

; TYPE: DNA

; ORGANISM: Homo Sapiens

US-11-266-748A-23628

Query Match 100.0%; Score 2831; DB 8; Length 2831;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2831; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCACAGCGGAGTTGGAGCGGATACGATTTGTGTGAGAGCGCAACGTGGCAT 60

DB 1 GGCACAGCGGAGTTGGAGCGGATACGATTTGTGTGAGAGCGGCAACGTGGCAT 60

QY 61 TCTGCTGACTTTGGAGGCATTTCTACGACTTTTCTCTCAGTGAGGCTTTTCTCCGACC 120

Db 61 |||||TCTGCTGAACTTGGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACC 120
QY 121 CTGATGCTCTTCAATTCGGTGTCTCCGACAGCCCAAGCTTGGCGTCTCTGAGAAATGGATGG 180
Db 121 CTGATGCTCTTCAATTCGGTGTCTCCGACAGCCCAAGCTTGGCGTCTCTGAGAAATGGATGG 180
QY 181 TCTTCACAAATAGCCCTTCAATTCGGTGTCTGAGTGGTATCAGTGCAGTGGTAAATGATGAA 240
Db 181 TCTTCACAAATAGCCCTTCAATTCGGTGTCTGAGTGGTATCAGTGCAGTGGTAAATGATGAA 240
QY 241 CACACTTCTTATGGAGAAACAGAGTCCAGTTCCTCTCTTTTGGATGTACCTTCTCTCTCT 300
Db 241 CACACTTCTTATGGAGAAACAGAGTCCAGTTCCTCTCTTTTGGATGTACCTTCTCTCTCT 300
QY 301 GCTCCCAATATGGAACATGTACTAGCAGTGGCAATGAAGAGGCTTTGTTCGATTTGAT 360
Db 301 GCTCCCAATATGGAACATGTACTAGCAGTGGCAATGAAGAGGCTTTGTTCGATTTGAT 360
QY 361 AACACAGAAATCACAAGTTTCAGAAAGAGTGCCTTCAAGAAATGGATGGCTCAGCTGGAAAT 420
Db 361 AACACAGAAATCACAAGTTTCAGAAAGAGTGCCTTCAAGAAATGGATGGCTCAGCTGGAAAT 420
QY 421 GCCGTCTTTGACCTGSCCTCGGGTTCCTGCTGAACTTAAACTTTGTTACAGCAGCAGTGTAT 480
Db 421 GCCGTCTTTGACCTGSCCTCGGGTTCCTGCTGAACTTAAACTTTGTTACAGCAGCAGTGTAT 480
QY 481 CAAACAGCCAAATTTTGGGACGTAAAGCTGGTGAGCTGATTTGGAACATGCAAAAGGTAT 540
Db 481 CAAACAGCCAAATTTTGGGACGTAAAGCTGGTGAGCTGATTTGGAACATGCAAAAGGTAT 540
QY 541 CAAATGAGCCTCAAGTCACTGCTTCTTCTTAAAGCTGAGTGCCTTCAAGAAAGTGGTAT 600
Db 541 CAAATGAGCCTCAAGTCACTGCTTCTTCTTAAAGCTGAGTGCCTTCAAGAAAGTGGTAT 600
QY 601 GGAAGAGATGGCAACATTTATGCTTGGGATACAGGTGCAACAAAGAAAGATGGTAT 660
Db 601 GGAAGAGATGGCAACATTTATGCTTGGGATACAGGTGCAACAAAGAAAGATGGTAT 660
QY 661 AGGCAAGTGAATCAAAATCAGTGGAGCTCAACATCTCAGACAAGCAAAACCCCTTCAAAA 720
Db 661 AGGCAAGTGAATCAAAATCAGTGGAGCTCAACATCTCAGACAAGCAAAACCCCTTCAAAA 720
QY 721 CCCAAGAGAAACAGAAATCAAAAGCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
Db 721 CCCAAGAGAAACAGAAATCAAAAGCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780
QY 781 ACTGTGGTCTCTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATA 840
Db 781 ACTGTGGTCTCTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATA 840
QY 841 ATCAAAGTATGGATTTTACGTAAAGAAATATATCTGCTTATCGACAAGAACCCCATAGCATCC 900
Db 841 ATCAAAGTATGGATTTTACGTAAAGAAATATATCTGCTTATCGACAAGAACCCCATAGCATCC 900
QY 901 AAGTCTTCTCTGATCCAGGTAGCAGCACTCGAAACCTTGGATATTAAGTCTGATTTTG 960
Db 901 AAGTCTTCTCTGATCCAGGTAGCAGCACTCGAAACCTTGGATATTAAGTCTGATTTTG 960
QY 961 GATTCACCTGGCTCTACTTTTATTTGCTAATTCACAGCAGATAACATCTACATGTTTAAAT 1020
Db 961 GATTCACCTGGCTCTACTTTTATTTGCTAATTCACAGCAGATAACATCTACATGTTTAAAT 1020
QY 1021 ATGACTGGGTGAAGACTTCTCAGTGGCTATTTTCAATGGACACCAAGAACTCTACCTTT 1080
Db 1021 ATGACTGGGTGAAGACTTCTCAGTGGCTATTTTCAATGGACACCAAGAACTCTACCTTT 1080
QY 1081 TATGTAAATCCAGCTTAGTCAGATGACAGTATTTAGTCAAGTGGCTCAAGTATGAA 1140
Db 1081 TATGTAAATCCAGCTTAGTCAGATGACAGTATTTAGTCAAGTGGCTCAAGTATGAA 1140
QY 1141 GCTGCTACATATGGAAGGTCTCCACACCTGGCAACCTTCTACTGTGCTCCCTGGGTAT 1200
Db 1141 GCTGCTACATATGGAAGGTCTCCACACCTGGCAACCTTCTACTGTGCTCTGCTGGGTAT 1200
QY 1201 TCTCAAGAGGTACGTCTGTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
Db 1201 TCTCAAGAGGTACGTCTGT 1260
QY 1261 TCTGATGACAAATACACTAAAAATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGGA 1320
Db 1261 TCTGATGACAAATACACTAAAAATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGGA 1320
QY 1321 GGTGATAAACTTTTCCACGGTGGGTGGGCTCTCAGAAAGAAAAAGAGTCAAGACCTGGC 1380
Db 1321 GGTGATAAACTTTTCCACGGTGGGTGGGCTCTCAGAAAGAAAAAGAGTCAAGACCTGGC 1380
QY 1381 CTAGTAACAGTAACGAGTAGCCAGAGTACTCTGCAAAAGCCCAAGGTTAAAGTGAAT 1440
Db 1381 CTAGTAACAGTAACGAGTAGCCAGAGTACTCTGCAAAAGCCCAAGGTTAAAGTGAAT 1440
QY 1441 CCATCCAAATTTTCCCGGTCTATCCGAGCTGTGCCCCCAAGCTGTCTGGAGACCTCCCT 1500
Db 1441 CCATCCAAATTTTCCCGGTCTATCCGAGCTGTGCCCCCAAGCTGTCTGGAGACCTCCCT 1500
QY 1501 CTTCTCTTCAAAATCTCTTACGTTCTCTATTTAAACCTCTCTGCAAGGCCCGGTCTCCC 1560
Db 1501 CTTCTCTTCAAAATCTCTTACGTTCTCTATTTAAACCTCTCTGCAAGGCCCGGTCTCCC 1560
QY 1561 ATCAACAGAGAGGCTGTCT 1620
Db 1561 ATCAACAGAGAGGCTGTCT 1620
QY 1621 TCGATTAGAAACTGGGTGACCCGCAACCTTCTCATCACCAACCTTCTCTCTCTCTCTCTCTCTCT 1680
Db 1621 TCGATTAGAAACTGGGTGACCCGCAACCTTCTCTCTCATCCACCCCTTCTCTCTCTCTCTCTCT 1680
QY 1681 TCGGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTTCTGTGAGCCAGAAAGTCACTCC 1740
Db 1681 TCGGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTTCTGTGAGCCAGAAAGTCACTCC 1740
QY 1741 CAAGCAGAGGCTTGTCTGTAGTCTAGAAATAGATTAAGAGAGGCTAGACTCAAGCTGT 1800
Db 1741 CAAGCAGAGGCTTGTCTGTAGTCTAGAAATAGATTAAGAGAGGCTAGACTCAAGCTGT 1800
QY 1801 CTGAGAGTGTGAAACAAAAGTGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGTATGGC 1860
Db 1801 CTGAGAGTGTGAAACAAAAGTGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGTATGGC 1860
QY 1861 CAAGTTGAAAAATCTTCAATTTGGATCTGTGCTGCCCTTGTGTTAAACCCAGAAAGACCTTAGT 1920
Db 1861 CAAGTTGAAAAATCTTCAATTTGGATCTGTGCTGCCCTTGTGTTAAACCCAGAAAGACCTTAGT 1920
QY 1921 AAGGACTCTTAGTCTTACCAATCAAGCAAAAATTGAAGAGGCTGGTACCAGTATCTCA 1980
Db 1921 AAGGACTCTTAGTCTTACCAATCAAGCAAAAATTGAAGAGGCTGGTACCAGTATCTCA 1980
QY 1981 GAGCTCCGTCTCTATCAGTCCGTGTGTTTCAAGAAAGCTGTGAAACCTTCTCTCTCTCTCTCT 2040
Db 1981 GAGCTCCGTCTCTATCAGTCCGTGTGTTTCAAGAAAGCTGTGAAACCTTCTCTCTCTCTCTCT 2040
QY 2041 TTGAGACCTTGTGGAGAGGCTCTGAAATGGTAGGCAAGAGAAATAGTTCCTCCAGAGAAAT 2100
Db 2041 TTGAGACCTTGTGGAGAGGCTCTGAAATGGTAGGCAAGAGAAATAGTTCCTCCAGAGAAAT 2100
QY 2101 AAAAATCTGGTGTGGCCATGGCAGCCAAACGGAAGGCTGAGAAATCCATCTCTCCAGAAAT 2160
Db 2101 AAAAATCTGGTGTGGCCATGGCAGCCAAACGGAAGGCTGAGAAATCCATCTCTCCAGAAAT 2160
QY 2161 CCGTCACTCCAGACACCAATTTCCAGGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTC 2220
Db 2161 CCGTCACTCCAGACACCAATTTCCAGGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTC 2220
QY 2221 ACCATACGCCCGAGCTCCATCAGGAAATCTGCATATCTTCCATAGAAAGTCTCCAGAG 2280
Db 2221 ACCATACGCCCGAGCTCCATCAGGAAATCTGCATATCTTCCATAGAAAGTCTCCAGAG 2280

QY 982 TTTGCTAATTGCACAGACGATAACATCTACATGTTTAAATATGACGTGGGTGGAAGCTTCT 1041
DB TTTGCTAATTGCACAGACGATAACATCTACATGTTTAAATATGACGTGGGTGGAAGCTTCT 1020
QY 1042 CCAGTGGCTATTTTCAATGGACACACAGAACTCTACCTTTTATGTAAATCCAGCCTTAGT 1101
DB CCAGTGGCTATTTTCAATGGACACACAGAACTCTACCTTTTATGTAAATCCAGCCTTAGT 1080
QY 1102 CCAGATGACCAAGTATTTAGTCAAGTGGCTCAAGTGAAGCTGCCTACATATGGAAGGTC 1161
DB CCAGATGACCAAGTATTTAGTCAAGTGGCTCAAGTGAAGCTGCCTACATATGGAAGGTC 1140
QY 1162 TCCACACCTCGGAACCTCTACTGTGCTCCTGGGTCAATCTCAAGAGGTCAAGTCTGTG 1221
DB TCCACACCTCGGAACCTCTACTGTGCTCCTGGGTCAATCTCAAGAGGTCAAGTCTGTG 1200
QY 1222 TGCTGGTGTCCATCTGACTTCAAAAGATTGCTACTGTTCTGTGATGACAATACACTAAAA 1281
DB TGCTGGTGTCCATCTGACTTCAAAAGATTGCTACTGTTCTGTGATGACAATACACTAAAA 1260
QY 1282 ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACACAGAGGTGATAAACTTTCCACGGTG 1341
DB ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACACAGAGGTGATAAACTTTCCACGGTG 1320
QY 1342 GGTGGGCTCTCAGAGAAAGAGTCAAGACCTGGCTAGTAACTAAAGAGTACGAGTAGC 1401
DB GGTGGGCTCTCAGAGAAAGAGTCAAGACCTGGCTAGTAACTAAAGAGTACGAGTAGC 1380
QY 1402 CAGAGTACTCTCCGCAAGCCCCCAGGGTAAAGTGCATCCATCAATCTTTCCCGCTCA 1461
DB CAGAGTACTCTCTCCGCAAGCCCCCAGGGTAAAGTGCATCCATCAATCTTTCCCGCTCA 1440
QY 1462 TCCGAGCTTGTGCCCCAAGCTGTGCTGGAGACCTCCCTCTTCTTCAAAATCTCCTACG 1521
DB TCCGAGCTTGTGCCCCAAGCTGTGCTGGAGACCTCCCTCTTCTTCAAAATCTCCTACG 1500
QY 1522 TTCTCTATTAAGACCTCTCTCCGAGGCCCGGTCTCCCATCAACAGAGAGGCTCTGTG 1581
DB TTCTCTATTAAGACCTCTCTCCGAGGCCCGGTCTCCCATCAACAGAGAGGCTCTGTG 1560
QY 1582 TCCTCGCTCTCCCAAGCCACTTCTCATCTTTCAAGATGTCCATTTAGAACTGGGTGACC 1641
DB TCCTCGCTCTCTCCCAAGCCACTTCTCATCTTTCAAGATGTCCATTTAGAACTGGGTGACC 1620
QY 1642 CGAACACCTTCTCATCACCACCCATCACTCCACCTGTCTCGGAGACCAAGATCATGTCT 1701
DB CGAACACCTTCTCATCACCACCCATCACTCCACCTGTCTCGGAGACCAAGATCATGTCT 1680
QY 1702 CCGAGAAAGCCCTTATTCCTGTGAGCCAGAAAGTCAATCCCAAGCAGAGCTTGTCTGAG 1761
DB CCGAGAAAGCCCTTATTCCTGTGAGCCAGAAAGTCAATCCCAAGCAGAGCTTGTCTGAG 1740
QY 1762 TCTAGAAATAGAGTAAAGAGGAGGTAGACTCAAGCTGTCTGGAGAGTGTGAACAAAG 1821
DB TCTAGAAATAGAGTAAAGAGGAGGTAGACTCAAGCTGTCTGGAGAGTGTGAACAAAG 1800
QY 1822 TGTGTGAAGAGTTGTAACCTGTGTGACTGAGCTTGTGTCGCAAGTGTGAAATCTTCATTTG 1881
DB TGTGTGAAGAGTTGTAACCTGTGTGACTGAGCTTGTGTCGCAAGTGTGAAATCTTCATTTG 1860
QY 1882 GATCTGTGCTGCTGTGCTGTAACCGAGAGCTTAGTAAGGACTCTTAGGTCTTAC 1941
DB GATCTGTGCTGCTGTGCTGTAACCGAGAGCTTAGTAAGGACTCTTAGGTCTTAC 1920
QY 1942 AAATCAAGCAAAATTCAGAGGCTGTACCACTACTCAGAGCTCCGCTCCTATCAGT 2001
DB AAATCAAGCAAAATTCAGAGGCTGTGTACCACTACTCAGAGCTCCGCTCCTATCAGT 1980
QY 2002 CCGTATGCTTCAGAAAGCTGTGGAAGCTTACCTCTTCTTTGAGACCTTTGTGGAGAGG 2061
DB CCGTATGCTTCAGAAAGCTGTGGAAGCTTACCTCTTCTTTGAGACCTTTGTGGAGAGG 2040

QY 2062 TCTCAATGCTAGGCAAGAGATAATGTTCCCGACAGATAAAACTGTTGTTGGCCATG 2121
DB TCTCAATGCTAGGCAAGAGATAATGTTCCCGACAGATAAAACTGTTGTTGGCCATG 2100
QY 2122 GCAGCCAAACGGAAGGCTGAGAATCCATCTCCAGAAAGTCCGTCTATCCAGACACCCAA 2181
DB GCAGCCAAACGGAAGGCTGAGAATCCATCTCCAGAAAGTCCGTCTATCCAGACACCCAA 2160
QY 2182 TCCAGGAGACAGAGCGGAAGACATTCGCAAGCCGGTCAACATCAAGCCAGCTCCATG 2241
DB TCCAGGAGACAGAGCGGAAGACATTCGCAAGCCGGTCAACATCAAGCCAGCTCCATG 2220
QY 2242 AGGAAATCTGCACATATCTCCATAGAAAGTCCAGAGGAGTCTCTGTGGTCCCTGAAAC 2301
DB AGGAAATCTGCACATATCTCCATAGAAAGTCCAGAGGAGTCTCTGTGGTCCCTGAAAC 2280
QY 2302 TCAACAGAAATATAGATTTCTAATCTGAGTGAATTTACTGAGCTTTGGTCCACTAAAA 2361
DB TCAACAGAAATATAGATTTCTAATCTGAGTGAATTTACTGAGCTTTGGTCCACTAAAA 2340
QY 2362 CTGAGCTTTGGTCCACTAAAAAAGATGAAAAATCAAGAGTACTCTATTAATCTCTGTC 2421
DB CTGAGCTTTGGTCCACTAAAAAAGATGAAAAATCAAGAGTACTCTATTAATCTCTGTC 2400
QY 2422 TTTAAGAAAGCTGCTTTTCAATTTTATAGACAAATCTTTTCAACGCTGAAATGTACTAA 2481
DB TTTAAGAAAGCTGCTTTTCAATTTTATAGACAAATCTTTTCAACGCTGAAATGTACTAA 2460
QY 2482 TCTGTTCTACTACATAATATATATGAGCTTCCCGAGGATGAATGCTGTTTAAAT 2541
DB TCTGTTCTACTACATAATATATATGAGCTTCCCGAGGATGAATGCTGTTTAAAT 2520
QY 2542 TCATAAAGTAAATTTGTCATCTCTAGCAATTTTGAATGAATAGTCTTCACTTTTAAAT 2601
DB TCATAAAGTAAATTTGTCATCTCTAGCAATTTTGAATGAATAGTCTTCACTTTTAAAT 2580
QY 2602 TCATCTCTCTATAATATGACATCCAGTTCATGGAGCAAAAAACAAGTTTCTTGTTA 2661
DB TCATCTCTCTATAATATGACATCCAGTTCATGGAGCAAAAAACAAGTTTCTTGTTA 2640
QY 2662 TCCTGAAACTTTCTATGCTCAGTGGAAAGTATCTGCCAGCCACAGCATGAGGCTGTGAA 2721
DB TCCTGAAACTTTCTATGCTCAGTGGAAAGTATCTGCCAGCCACAGCATGAGGCTGTGAA 2700
QY 2722 GGCTGACTGAGAAATCTCTCTGCAAGACCCCTGGTCTGTGCTTCCCAATGTATA 2781
DB GGCTGACTGAGAAATCTCTCTGCAAGACCCCTGGTCTGTGCTTCCCAATGTATA 2760
QY 2782 ATTTTATTTGAATACATAATCTTTTCACTATG 2814
DB ATTTTATTTGAATACATAATCTTTTCACTATG 2793

RESULT 3
US-11-266-748A-56960
; Sequence 56960, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55915-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266.748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0

```

; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 493996
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 56960
; LENGTH: 4221
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-749A-56960

```

Query Match 98.5%; Score 2789.8; DB 8; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	22	CGATAACGATTTGTGTGTGAGAGCGCGAACGTCGGAGTTTCTGCTGAACTTTGGAGGCATTT	81
DB	1	CGATAACGATTTGTGTGTGAGAGCGCGAACGTCGATTTTCTGCTGAACTTTGGAGGCATTT	60
QY	82	TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCTCGACCCCTGATGCTCTTCAATTTCGGTG	141
DB	61	TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCTCGACCCCTGATGCTCTTCAATTTCGGTG	120
QY	142	CTCGCCAGCCGCCACGCTTGGCGTCTCTGAGAAATGGATGGTCTTTCACAAATACCCTCTTCAA	201
DB	121	CTCGCCAGCCGCCACGCTTGGCGTCTCTGAGAAATGGATGGTCTTTCACAAATACCCTCTTCAA	180
QY	202	TCCCTTCTGACTGGTTATCAGTCAGATGGTGAATGATGAACAACACTTCTTATGGAGAAACA	261
DB	181	TCCCTTCTGACTGGTTATCAGTCAGATGGTGAATGATGAACAACACTTCTTATGGAGAAACA	240
QY	262	GGAGTCCAGTTCCTCTCTTTTGGATGTACCTTCTCTCTGCTCCCAATATGGAACATGTA	321
DB	241	GGAGTCCAGTTCCTCTCTTTTGGATGTACCTTCTCTCTGCTCCCAATATGGAACATGTA	300
QY	322	CTAGCAGTTGCCAATGAAGAAGCTTTGTTCGATTTGTATATAACAGAAATCACAAAGTTTC	381
DB	301	CTAGCAGTTGCCAATGAAGAAGCTTTGTTCGATTTGTATATAACAGAAATCACAAAGTTTC	360
QY	382	AGAAAGAGTGTCTTCAAGAATGATGGCTCAGCTGGAATGCGCTTTTGACCTGGCCCTGG	441
DB	361	AGAAAGAGTGTCTTCAAGAATGATGGCTCAGCTGGAATGCGCTTTTGACCTGGCCCTGG	420
QY	442	GTTCCTGCTGAACCTTAAACTTTGTACAGCAGCAGGTGATCAAAACAGCCAAATTTTGGGAC	501
DB	421	GTTCCTGCTGAACCTTAAACTTTGTACAGCAGCAGGTGATCAAAACAGCCAAATTTTGGGAC	480
QY	502	GTAAGAGCTGGTGAGCTGATTTGGAACATGCAAGGTCATCAATGCAGCCTCAAGTCAGTT	561
DB	481	GTAAGAGCTGGTGAGCTGATTTGGAACATGCAAGGTCATCAATGCAGCCTCAAGTCAGTT	540
QY	562	GCCTTTTCTAAGTTTGAAGAGCTGTATTCTGTACGGGTGGAAGAGATGGCAACATTTATG	621
DB	541	GCCTTTTCTAAGTTTGAAGAGCTGTATTCTGTACGGGTGGAAGAGATGGCAACATTTATG	600
QY	622	GTCTGGGATACCCAGGTGCACAAAAAAGATGGGTTTTATAGGCAAGTCAATCAAAATCAGT	681
DB	601	GTCTGGGATACCCAGGTGCACAAAAAAGATGGGTTTTATAGGCAAGTGAATCNAATCAGT	660
QY	682	GGAGCTCACAATACCTCAGACAAGCAAAACCCCTTCCAAACCCAGAGAGAAACAGAAATTC	741
DB	661	GGAGCTCACAATACCTCAGACAAGCAAAACCCCTTCCAAACCCAGAGAGAAACAGAAATTC	720
QY	742	AAAGGACTTGTCTCTCTGTGGATTTCCAGCAAAAGTGTATCTGTGGTCTCTTCTCAAGAC	801
DB	721	AAAGGACTTGTCTCTCTGTGGATTTCCAGCAAAAGTGTATCTGTGGTCTCTTCTCAAGAC	780

			C A A A G T A T G G G A T T A C G T A A G A A T T A T A C T G C T T A T T C G A C A A G A C C C A T A G C A T C C A A	765
Db				
Qy	903	G T C T T T C T G T A C C C A G G T A G C A G C A C T C G A A A A C T T T G G A T A T T C A A G T C T G A T T T T T G G A	962	
Dd	766	G T C T T T C T G T A C C C A G G T A G C A G C A C T C G A A A A C T T T G G A T A T T C A A G T C T G A T T T T T G G A	825	
Qy	963	T T C A C T G G C T A C T A T T A T T T G C T A A T T G C A C A G A C G A T A A C A T C T A C A T G T T T A A T A T	1022	
Dd	826	T T C A C T G G C T A C T A T T A T T T G C T A A T T G C A C A G A C G A T A A C A T C T A C A T G T T T A A T A T	885	
Qy	1023	G A C T G G G T T G N A G A C T T C T C A G T G G C T A T T T T C A A T G E A C A C A G A A C T C T A C C I T T T A	1082	
Dd	886	G A C T G G G T T G A A G A C T T C T C C A G T G G C T A T T T T C A A T G G A C A C A G A A C T C T A C C I T T T A	945	
Qy	1083	T G T A A A A T C C A G C C T T A G T C C A G A T A C C A G T T T T T A G T C A G T G G C T C A A G T G A T G A A G C	1142	
Dd	946	T G T A A A A T C C A G C C T T A G T C C A G A T A C C A G T T T T T A G T C A G T G G C T C A A G T G A T G A A G C	1005	
Qy	1143	T G C C T A C A T A T G S A A G G T C T C C A C A C C C T G G C A A C C T C C T A C T G T G C T C C T G G G T C A A T T C	1202	
Dd	1006	T G C C T A C A T A T G S A A G G T C T C C A C A C C C T G G C A A C C T C C T A C T G T G C T C C T G G G T C A A T T C	1065	
Qy	1203	T C A A G A G G T C A C G T C T G T G C T G G T C C A T C T G A C T T C A A A A G A T T G C T A C C T G T T C	1262	
Dd	1066	T C A A G A G G T C A C G T C T G T G T G C T G G T G C C A T C T G A C T T C A A A A G A T T G C T A C C T G T T C	1125	
Qy	1263	T G A T G C A A T A C A C T A A A A A T C T G G C G C T T G S A A T A G A G G C T T A G A G A G A A A C C A G A G G	1322	
Dd	1126	T G A T G C A A T A C A C T A A A A T C T G G C G T T G A A T A G A G G C T T A G A G A G A A A C C A G A G G	1185	
Qy	1323	T G A T A A A C T T T T C C A C G G T G G G T T G G G C C T C T C A G A A G A A A A A A A G A G T C A A G A C C T G G C C T	1382	
Dd	1186	T G A T A A A C T T T T C C A C G G T G G G T T G G G C C T C T C A G A A G A A A A A A A G A G T C A A G A C C T G G C C T	1245	
Qy	1383	A G T A A C A G T A A C A G A G T A G C C A G A G T A C T C T G C C A A A G C C C C A G G G T A A A G T G C A A T T C C	1442	
Dd	1246	A G T A A C A G T A A C A G A G T A G C C A G A G T A C T C T G C C A A A G C C C C A G G G T A A A G T G C A A T T C C	1305	
Qy	1443	A T C C A A T T C T C C C G G T C A T C C G C A G C T T G T G C C C A A G C T G T G C T C G A G A C C C T C C C T C T	1502	
Dd	1306	A T C C A A T T C T C C C C G T C A T C C G C A G C T T G T G C C C C A A G C T G T G C T G G A G A C C T C C C T C T	1365	
Qy	1503	T C C T T C A A A T A C T C C T A C G T T C T C T A T T A A A A C C T C T C T G C C A A G C C C G G T C T C C C A T	1562	
Dd	1366	T C C T T C A A A T A C T C C T A C G T T C T C T A T T A A A A C C T C T C C T G C C A A G C C C G G T C T C C C A T	1425	
Qy	1563	C A A C A G A A G G C T C T G T C C T C C G T C T C T C C A A G C A C C T T C A T C T T T C A A G A T G T C	1622	
Dd	1426	C A A C A G A A G G C T C T G T C T C C T C C G T C T C T C C A A G C C A C C T T C A T C T T T C A A G A T G T C	1485	
Qy	1623	G A T T A G A A A C T G G G T G A C C G A A C A C C T T C C T C A T C A C C A C C C A T C A C C A C C T G C T C	1682	
Dd	1486	G A T T A G A A A C T G G G T G A C C C G A A C A C C T T C C T A T C A C C A C C C A T C A C C A C C T G C T C	1545	
Qy	1683	G G A C C A A A G A T C A T G T C T C G A A A A G C C C T T A T T C C T G T G A G C C A G A A G T C A T C C C A	1742	
Dd	1546	G G A C C A A A G A T C A T G T C T C G A A A A A A G C C C T T A T T C C T G T G A G C C A G A A G T C A T C C C A	1605	
Qy	1743	A G C A G A G C T T G C T C T G A G T C T A G A A A T A G A G T A A A G A G A G G C T A G A C T C A A G C T G T C T	1802	
Dd	1606	A G C A G A G C T T G C T C T G A G T C T A G A A A T A G A G T A A A G A G A G G C T A G A C T C A A G C T G T C T	1665	
Qy	1803	G G A G A G T G A A A C A A A A G T G T G A A G A G T T G T A A C T G T G T A C T A G A C T T G A T G G C C A	1862	
Dd	1666	G G A G A G T G A A A C A A A A G T G T G A A G A G T T G T A A C T G T G T A C T A G A C T T G A T G G C C A	1725	
Qy	1863	A G T T G A A A A C T C T C A T T T T G G A T C T G T G C T G C T T G T G T A A C C A G A A A C C T T A G T A A	1922	
Dd	1726	A G T T G A A A A C T C T C A T T T T G G A T C T G T G C T G C T T G C T G G T A A C C A G A A A C C T T A G T A A	1785	
Qy	1923	G G A C T C T C T A G G T C C T A C C A A A T C A A G C A A A T T G A A G A G A G C T G G T A C C A G A T A T C T C A G A	1982	

Db	1786	GGA	CTCTCTAGGTCCTACCAAAATCAACAAAAATTGAAGGAGCTGGTACCAGTATCTCAGA	1845
Qy	1983	GCCTCCGTCCTCATCAGTCGGTAGCTTCAGAAAAGCTGTGAAACGTCCTCTTCCTTT	2042	
Db	1846	GCCTCCGTCCTCATCAGTCGGTAGCTTCAGAAAAGCTGTGAAACGTCCTCTTCCTTT	1905	
Qy	2043	GAGACCTTGCGAAGGGCTCGAAATGGTTAGCCNAAGANAATAGTTCCGCCAGANATAA	2102	
Db	1906	GAGACCTTGCGAAGGGCTCGAAATGGTAGGCAAGAGAAATAGTTCCGCCAGANATAA	1965	
Qy	2103	AAACTGGTTGTTGGCCATGGCAGCCAACGGAAGGCTGAGAATCCATCTCCACGAAGTCC	2162	
Db	1966	AAACTGGTTGTTGGCCATGGCAGCCAACGGAAGGCTGAGNATCCATCTCCACGAAGTCC	2025	
Qy	2163	GTCATCCGAGACACCCAAATTCAGGAGACAGAGCGGAAGAACATTTGCCAAGCCCGTTCAC	2222	
Db	2026	GTCATCCGAGACACCCAAATTCAGGAGACAGAGCGGAAGAAATTTGCCAAGCCCGTTCAC	2085	
Qy	2223	CATCACGCCCAGCTCCATGAGGAAAAATCTGCACATCTTCATAGAAGTCCCAGGAGGA	2282	
Db	2086	CATCACGCCCAGCTCCATGAGGAAAAATCTGCACATCTTCATAGAAGTCCCAGGAGGA	2145	
Qy	2283	CTTCTGTGGTCTGAAACACTCAACAGCAAATTTATAGATTTCTAATCTGAGTGAGTTACTGAGC	2342	
Db	2146	CTTCTGTGGTCTGAAACACTCAACAGCAAATTTATAGATTTCTAATCTGAGTGAGTTACTGAGC	2205	
Qy	2343	TTTTGGTCCATAAAACAAGCTGAGCTTTGGTCCACTAAACACAGATGAAAATACAGAG	2402	
Db	2206	TTTTGGTCCATAAAACAAGCTGAGCTTTGGTCCACTAAACACAGATGAAAATACAGAG	2365	
Qy	2403	TGACTCTATAACTCTGGTCTTTTAAGAAGCTGGCTTTTCATTTTATAGACAAAATCTTTTC	2462	
Db	2266	TGACTCTATAACTCTGGTCTTTTAAGAAGCTGGCTTTTCATTTTATAGACAAAATCTTTTC	2325	
Qy	2463	AACGCTGAAATGTACCTAAATCTGTTCTCTACATAAATGTATATGACAGCTTCCCGAGGA	2522	
Db	2326	AACGCTGAAATGTACCTAAATCTGTTCTCTACATAAATGTATATGACAGCTTCCCGAGGA	2385	
Qy	2523	TGAATGCTGTGTTTAAATTTTCATAAAGTAAATTTGTGCACCTAGCATTTTGAATGAATAG	2582	
Db	2386	TGAATGCTGTGTTTAAATTTTCATAAAGTAAATTTGTGCACCTAGCATTTTGAATGAATAG	2445	
Qy	2583	TCITTCACTTTTAAATTAATTCATCTCTCTATAATAATGACATCCCCAGTTTCATGAGGCA	2642	
Db	2446	TCITTCACTTTTAAATTAATTCATCTCTCTATAATAATGACATCCCCAGTTTCATGAGGCA	2505	
Qy	2643	AAAAACAAGTTTCTTGTTATCTCGTGAACCTTTCTATGCTCAGTGGAAAGTATCTGCCAGCC	2702	
Db	2506	AAAAACAAGTTTCTTGTTATCTCGTGAACCTTTCTATGCTCAGTGGAAAGTATCTGCCAGCC	2565	
Qy	2703	ACAGCATGAGGCCCTGTGAAGGCTGACTGAGAAATCCCTCTGCTGAAGACCCCTTGTTCTGT	2762	
Db	2566	ACAGCATGAGGCCCTGTGAAGGCTGACTGAGAAATCCCTCTGCTGAAGACCCCTTGTTCTGT	2625	
Qy	2763	TC TGCCCTCCAACTGATAAATTTTATTTGAAAATACATAAATCTTTTCTACTATG	2814	
Db	2626	TC TGCCCTCCAACTGATAAATTTTATTTGAAAATACATAAATCTTTTCTACTATG	2677	

```

RESULT 5
US-11-2666-748A-184890
; Sequence 184890, Application US/112666748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE OF INVENTION: Methods of Using the Same
; TITLE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266, 748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2

```



```

; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 48396
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 184890
; LENGTH: 1949
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-184890

Query Match      58.4%; Score 1937.4; DB 8; Length 1949;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1941; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 12 GAGTTGGAGGCGATACGATTTGTGTGTGAGAGGCGCAACGCTGGGATTTCTGCTGAAC 71
DB 3 GGGTTGGAGGCGATACGATTTGTGTGTGAGAGGCGCAAGCTGCGATTTCTGCTGAAC 62

QY 72 TGGAGGCAATTTCTACGACTTTTCTCTCAGCTAGGCTTTTCTCGACCCCTGATGCTCT 131
DB 63 TGGAGGCAATTTCTCAGCTAGGCTTTTCTCAGCTAGGCTTTTCTCGACCCCTGATGCTCT 122

QY 132 CAATTCGGTGCTCCGCGACGCCAGCTTGGCGTCCCTGAGAAATGGATGCTTTCAACAATA 191
DB 123 CAATTCGGTGCTCCGCGACGCCAGCTTGGCGTCCCTGAGAAATGGATGCTTTCAACAATA 182

QY 192 CCCTCTTCAATCCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACAACATCTCTTA 251
DB 183 CCCTCTTCAATCCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACAACATCTCTTA 242

QY 252 TGGAGAAACAGAGTCCAGTTCCTCTCTTTGGATGTACCTTCTCTCTGCTCCCAATAT 311
DB 243 TGGAGAAACAGAGTCCAGTTCCTCTCTTTGGATGTACCTTCTCTCTGCTCCCAATAT 302

QY 312 GGAACATGTACTAGCAGTTGCCAATGAAGAAGCTTTGTTGATTTGATTAACACAGAATC 371
DB 303 GGAACATGTACTAGCAGTTGCCAATGAAGAAGCTTTGTTGATTTGATTAACACAGAATC 362

QY 372 ACAAGTTTTCAGAAAGAGTGCTTCAAGAAATGGATGGCTCACTGGAATGCCGCTTTTGA 431
DB 363 ACAAGTTTTCAGAAAGAGTGCTTCAAGAAATGGATGGCTCACTGGAATGCCGCTTTTGA 422

QY 432 CTGCGCTGGGTTCTGTTGAACTTAAACTTTGTTACAGCAGCAGGTGATCAAAAGCCAA 491
DB 423 CTGCGCTGGGTTCTGTTGAACTTAAACTTTGTTACAGCAGCAGGTGATCAAAAGCCAA 482

QY 492 ATTTTGGGACGTAAAGCTGGTGAGCTGATTTGGAACATGCAAGGTTCATCAATGCAGCT 551
DB 483 ATTTTGGGACGTAAAGCTGGTGAGCTGATTTGGAACATGCAAGGTTCATCAATGCAGCT 542

QY 552 CAAGTCAGTTGCTTTTCTTAAGTTTGGAGAAAGCTGTATTTCTGTACGGGTGGAAGAGATGG 611
DB 543 CAAGTCAGTTGCTTTTCTTAAGTTTGGAGAAAGCTGTATTTCTGTACGGGTGGAAGAGATGG 602

QY 612 CAACATTATGCTTGGGATACCAAGTGGCAACAAAAAGATGGGTTTATAGGCAAGTGAA 671
DB 603 CAACATTATGCTTGGGATACCAAGTGGCAACAAAAAGATGGGTTTATAGGCAAGTGAA 662

QY 672 TCAATTCAGTGGGCTCAATACCTCAGCAAGCAACCCCTTCAAAACCCCAAGAGAA 731
DB 672 TCAATTCAGTGGGCTCAATACCTCAGCAAGCAACCCCTTCAAAACCCCAAGAGAA 731

```

```

DB 663 TCAATTCAGTGGGCTCAATACCTCAGCAAGCAACCCCTTCAAAACCCCAAGAGAA 722
QY 732 ACAGAAATTCAGAAAGAGCTTGTCTCTCTGTGGATTTTCAGCAAAAGTGTACTGTGGTCT 791
DB 723 ACAGAAATTCAGAAAGAGCTTGTCTCTCTGTGGATTTTCAGCAAAAGTGTACTGTGGTCT 782
QY 792 CTTTTCAGAGCAGAGATACCTTAGTCTCAGCAGGAGCTGTGTGATGGGATAAATCAAAAGTATG 851
DB 783 CTTTTCAGAGCAGAGATACCTTAGTCTCAGCAGGAGCTGTGTGATGGGATAAATCAAAAGTATG 842
QY 852 GGATTTTACGTAAGATTAATATCTGCTTATTCGACAAGAACCCATAGCATCCAAAGTCTTTTCT 911
DB 843 GGATTTTACGTAAGATTAATATCTGCTTATTCGACAAGAACCCATAGCATCCAAAGTCTTTTCT 902
QY 912 GTACCCAGGTAGCAGCACTCGAAATCTTGATATTTCAAGTCTGATTTGGATTTCCACATGG 971
DB 903 GTACCCAGGTAGCAGCACTCGAAATCTTGATATTTCAAGTCTGATTTGGATTTCCACATGG 962
QY 972 CTCTACTTTTATTTGCTAAATTCGACAGAGATAAATCTACATGTTTAAATATGACTGGGTT 1031
DB 963 CTCTACTTTTATTTGCTAAATTCGACAGAGATAAATCTACATGTTTAAATATGACTGGGTT 1022
QY 1032 GAAGACTTCTCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTATGTAAAAATC 1091
DB 1023 GAAGACTTCTCAGTGGCTATTTTCAATGGTGGCCAGAACTCTACCTTTTATGTAAAAATC 1082
QY 1092 CAGCTTTAGTCCAGATGACCAAGTTTTAGTCAAGTGGCTCAAGTGATGATGATGCTGCTACAT 1151
DB 1083 CAGCTTTAGTCCAGATGACCAAGTTTTAGTCAAGTGGCTCAAGTGATGATGATGCTGCTACAT 1142
QY 1152 ATGGAAGTCTCCACACCTCGCAACCTCTTACTGTGCTCTGCTGGTCAATCTCAAGAGGT 1211
DB 1143 ATGGAAGTCTCCACACCTCGCAACCTCTTACTGTGCTCTGCTGGTCAATCTCAAGAGGT 1202
QY 1212 CAGCTCTGTGCTGCTGCTCCATCTGACTTCAACAAGATTGCTTACTGTTCTGTGATGACAA 1271
DB 1203 CAGCTCTGTGCTGCTGCTCCATCTGACTTCAACAAGATTGCTTACTGTTCTGTGATGACAA 1262
QY 1272 TACACTAAAAATCTGGCGCTTGAATAGAGGCTTAGAGAGAGAAACAGAGAGGTGATAAATC 1331
DB 1263 TACACTAAAAATCTGGCGCTTGAATAGAGGCTTAGAGAGAGAAACAGAGAGGTGATAAATC 1322
QY 1332 TTCCACGGTGGGTTGGGCTCTCAGAGAGAAAGAGTCAAGACTGSCCTAGTAACAGT 1391
DB 1323 TTCCACGGTGGGTTGGGCTCTCAGAGAGAAAGAGTCAAGACTGSCCTAGTAACAGT 1382
QY 1392 AACGAGTAGCCAGAGTACTCCTGCAAGAGCCCGCAGGGTAAAGTGCATCCATCCAAATC 1451
DB 1383 AACGAGTAGCCAGAGTACTCCTGCAAGAGCCCGCAGGGTAAAGTGCATCCATCCAAATC 1442
QY 1452 TTCCCGGTCATCCGAGCTTGTGCCCCCAAGCTGTGTGCGAGACCTCCCTCTTCTTCAAA 1511
DB 1443 TTCCCGGTCATCCGAGCTTGTGCCCCCAAGCTGTGTGCGAGACCTCCCTCTTCTTCAAA 1502
QY 1512 TACTCTTACGTTCTCTATTAAGCTCTCTGCAAGCCCGGCTCTCCCATCAACAGAG 1571
DB 1503 TACTCTTACGTTCTCTATTAAGCTCTCTGCAAGCCCGGCTCTCCCATCAACAGAG 1562
QY 1572 AGGCTCTGCTCTCTCCGCTCTCTCCCAAGCCACCTTTCATCTTTCAAGATGTCGATAGAAA 1631
DB 1563 AGGCTCTGCTCTCTCCGCTCTCTCCCAAGCCACCTTTCATCTTTCAAGATGTCGATAGAAA 1622
QY 1632 CTGGGTGACCCGAAACACCTTCTCTATCAACCCATCACTCCACTGCTTGGAGACCA 1691
DB 1623 CTGGGTGACCCGAAACACCTTCTCTATCAACCCATCACTCCACTGCTTGGAGACCA 1682
QY 1692 GATCATGTCTCCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTGCATCCCAAGCAGAGGC 1751
DB 1683 GATCATGTCTCCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTGCATCCCAAGCAGAGGC 1742
QY 1752 TTGCTCTGAGTCTAGAAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCTGGAGAGTGT 1811
DB 1743 TTGCTCTGAGTCTAGAAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCTGGAGAGTGT 1802

```


Query Match	43.6%;	Score 1235;	DB 8;	Length 1270;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 1246;	Conservative 0;	Mismatches 0;	Indels 1;	Gaps 1;
QY	1202	CTCAAGAGGTCACGCTCTGTGTGCTGGTGTCCATCTGACTTCACAAAGATTGCTACCTGTT	1261	
DB	24	CTCAAGAGGTCACGCTCTGTGTGCTGGTGTCCATCTGACTTCACAAAGATTGCTACCTGTT	83	
QY	1262	CTGATGACAATACACTAAATCTGGCCGCTTGAATAGAGGCTTAGAGGAGAAACACGAG	1321	
DB	84	CTGATGACAATACACTAAATCTGGCCGCTTGAATAGAGGCTTAGAGGAGAAACACGAG	143	
QY	1322	GTGATAAATCTTTCCACGGTGGGTTGGGCTCTCTCAGAAGAAAAAGAGTCAAGACCTGGCC	1381	
DB	144	GTGATAAATCTTTCCACGGTGGGTTGGGCTCTCTCAGAAGAAAAAGAGTCAAGACCTGGCC	203	
QY	1382	TAGTAAACAGTAAACGAGTAGCCAGAGTACTCTCTGCCAAGCCCCCAGGGTAAAGTGCAATC	1441	
DB	204	TAGTAAACAGTAAACGAGTAGCCAGAGTACTCTCTGCCAAGCCCCCAGGGTAAAGTGCAATC	263	
QY	1442	CATCCAATCTTCCCGCTCATCCGAGCTTGCCGCCAAGCTGTCTGGAGACCTCCCTC	1501	
DB	264	CATCCAATCTTCCCGCTCATCCGAGCTTGCCGCCAAGCTGTCTGGAGACCTCCCTC	323	
QY	1502	TTTCCTTCAAATACTCTACGTTCTCTATTAATAAACCCTCTCTGCCAAGCCCCGGTCTCCCA	1561	
DB	324	TTTCCTTCAAATACTCTACGTTCTCTATTAATAAACCCTCTCTGCCAAGCCCCGGTCTCCCA	383	
QY	1562	TCAACAGAGAGGCTCTGTCTCTCCGCTCTCTCCCAAGCCACCTTCATCTTTCAAGATGT	1621	
DB	384	TCAACAGAGAGGCTCTGTCTCTCCGCTCTCTCCCAAGCCACCTTCATCTTTCAAGATGT	443	
QY	1622	CGATTAGAACTGGGTGACCCGAAACCTCTCTATCAACACCACTCACTCCACCTGTT	1681	
DB	444	CGATTAGAACTGGGTGACCCGAAACCTCTCTATCAACACCACTCACTCCACCTGTT	503	
QY	1682	CGGAGACCAAGTCACTGTCTCGGAGAAAAAGCCCTTATTCCTGTGAGCCAGAGTCACTCC	1741	
DB	504	CGGAGACCAAGTCACTGTCTCGGAGAAAAAGCCCTTATTCCTGTGAGCCAGAGTCACTCC	563	
QY	1742	TAGCAGAGGCTTGCTCTGAGTCTAGAAATAGAGTAAAGAGAGGCTAGACTCAAGCTGTC	1801	
DB	564	TAGCAGAGGCTTGCTCTGAGTCTAGAAATAGAGTAAAGAGAGGCTAGACTCAAGCTGTC	623	
QY	1802	TGGAGAGTGTAAACAAAGTGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGATGCC	1861	
DB	624	TGGAGAGTGTAAACAAAGTGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGATGCC	683	
QY	1862	AAGTTGAAAACTTTCATTGGATCTGTCTGCCCTTGTGGTAAACAGGAACCTTAGTA	1921	

Db	768	AAATCCTCTGCTGAAGACCCCTGGTTCTGTCTGCTCCCAACATGTATTAATTTATTGA	827
QY	2793	ATAACATAATCTTTTCACTATG	2814
Db	828	ATAACATAATCTTTTCACTATG	849
RESULT 11			
US-11-266-748A-146319/c			
; Sequence 146319, Application US/11266748A			
; Publication No. US20060134663A1			
; GENERAL INFORMATION:			
; APPLICANT: Harkin, Paul			
; APPLICANT: Johnston, Patrick			
; APPLICANT: Mulligan, Karl			
; TITLE OF INVENTION: Transcriptome Microarray Technology and			
; FILE REFERENCE: 55815-0102 (319189)			
; CURRENT APPLICATION NUMBER: US/11/266,748A			
; CURRENT FILING DATE: 2005-11-03			
; PRIOR APPLICATION NUMBER: EP 04105479.2			
; PRIOR FILING DATE: 2004-11-03			
; PRIOR APPLICATION NUMBER: EP 04105482.6			
; PRIOR FILING DATE: 2004-11-03			
; PRIOR APPLICATION NUMBER: EP 04105483.4			
; PRIOR FILING DATE: 2004-11-03			
; PRIOR APPLICATION NUMBER: EP 04105507.0			
; PRIOR FILING DATE: 2004-11-03			
; PRIOR APPLICATION NUMBER: EP 04105485.9			
; PRIOR FILING DATE: 2004-11-03			
; PRIOR APPLICATION NUMBER: EP 04105484.2			
; PRIOR FILING DATE: 2004-11-03			
; PRIOR APPLICATION NUMBER: US 60/662,276			
; PRIOR FILING DATE: 2005-03-14			
; PRIOR APPLICATION NUMBER: US 60/700,293			
; PRIOR FILING DATE: 2005-07-18			
; NUMBER OF SEQ ID NOS: 483996			
; SOFTWARE: PatentIn version 3.3			
; SEQ ID NO 146319			
; LENGTH: 931			
; TYPE: DNA			
; ORGANISM: Homo Sapiens			
; FEATURE:			
; NAME/KEY: misc_feature			
; LOCATION: (62)..(71)			
; OTHER INFORMATION: n is a, c, g, or t			
US-11-266-748A-146319			
Query Match			
Best Local Similarity 28.3%; Score 802; DB 8; Length 931;			
Matches 802; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	2013	AGAAAGCTGTGAACGCTACCTCTTCTTTGAGACCTCTGTGGAGAGGCTGTGAATGTG	2072
Db	884	AGAAAGCTGTGAACGCTACCTCTTCTTTGAGACCTCTGTGGAGAGGCTGTGAATGTG	825
QY	2073	AGGCAAGAGATAGTTCCTCCAGAGATATAAATCTGGTTGGCCATGGCAGCCAAAGC	2132
Db	824	AGGCAAGAGATAGTTCCTCCAGAGATATAAATCTGGTTGGCCATGGCAGCCAAAGC	765
QY	2133	GAAGGCTGAGATCCATCTCCAGAGTCCGTCATCCAGACACCAATTCAGGAGACA	2192
Db	764	GAAGGCTGAGATCCATCTCCAGAGTCCGTCATCCAGACACCAATTCAGGAGACA	705
QY	2193	GAGCGAAGACATTCGCAAGCCCGGTACCATCCAGCCAGCTCCATCAGGAAATCTG	2252
Db	704	GAGCGAAGACATTCGCAAGCCCGGTACCATCCAGCCAGCTCCATCAGGAAATCTG	645
QY	2253	CACATATCTCCATAGAAAGTCCAGAGGACTTCTGTGTCTCTGACACTCAACAGAAAT	2312
Db	644	CACATATCTCCATAGAAAGTCCAGAGGACTTCTGTGTCTCTGACACTCAACAGAAAT	585

RESULT 12
US-11-266-748A-41156
; Sequence 41156, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 41156
; LENGTH: 883
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-11-266-748A-41156


```
QY 2674 CTATGCTCAGTGAAGATCTGCCAGCCACAGCATGAGGCCCTGTGAAGGCTGACTGAGA 2733
Db 661 CTATGCTCAGTGAAGATCTGCCAGCCACAGCATGAGGCCCTGTGAAGGCTGACTGAGA 720
QY 2734 AATCCTCTGCTGAAGACCCCTGGTCTGTGTTCTGCCCTCCAAATGATATAATTTATTTGAA 2793
Db 721 AATCCTCTGCTGAAGACCCCTGGTCTGTGTTCTGCCCTCCAAATGATATAATTTATTTGAA 780
QY 2794 ATACATAAATCTTTTCACTATG 2814
Db 781 ATACATAAATCTTTTCACTATG 801

RESULT 14
US-11-266-748A-261935
; Sequence 261935, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 60/662,276
; PRIOR FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/700,293
; PRIOR FILING DATE: 2005-07-18
; NUMBER OF SEQ ID NOS: 483996
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 261935
; LENGTH: 798
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (706)..(706)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (728)..(728)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (760)..(760)
; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-261935

Query Match 24.8%; Score 700.8; DB 8; Length 798;
Best Local Similarity 97.0%; Pred. No. 2.5e-185;
Matches 756; Conservative 0; Mismatches 15; Indels 8; Gaps 4;

QY 795 TCAAGACGAGATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATATCAAGATATGGA 854
Db 14 TCAAGACGAGATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATATCAAGATATGGA 73
QY 855 TTATACCTAAGAATTATCTGCTTATCGACAAGAACCATAGCATCCAAGTCTTTCTCTGTA 914
```

```
Db 74 TTATACCTAAGAATTATCTGCTTATCGACAAGAACCATAGCATCCAAGTCTTTCTCTGTA 133
QY 915 CCAGGTAGCAGCAGCTCGAATACTTTGGATATTCAAGTCTGATTTTGGATTTCCACTGGCTC 974
Db 134 CCAGGTAGCAGCAGCTCGAATACTTTGGATATTCAAGTCTGATTTTGGATTTCCACTGGCTC 193
QY 975 TACTTTATTGCTAAATTTGCACAGAGATAACATCTACATGTTTAAATATGACTGGGTTGAA 1034
Db 194 TACTTTATTGCTAAATTTGCACAGAGATAACATCTACATGTTTAAATATGACTGGGTTGAA 253
QY 1035 GACTTCTCAGTGGCTATTTTCAATGGACACAGAACTCTACTCTTTATGTAATCCAG 1094
Db 254 GACTTCTCAGTGGCTATTTTCAATGGACACAGAACTCTACTCTTTATGTAATCCAG 313
QY 1095 CCTTAGTCCAGATGACCAAGTCTTTTGTAGTCAAGTGGCTCAAGTATGAAGTCCCTACATAG 1154
Db 314 CCTTAGTCCAGATGACCAAGTCTTTTGTAGTCAAGTGGCTCAAGTATGAAGTCCCTACATAG 373
QY 1155 GAAGGTCTCCACACCCCTGGCAACCTCTACTGTGCTCCTGGGTCAATTTCTCAAGAGGTCA 1214
Db 374 GAAGGTCTCCACACCCCTGGCAACCTCTACTGTGCTCCTGGGTCAATTTCTCAAGAGGTCA 433
QY 1215 GTCTGTGCTGGTGTCCATCTGACTTCAAAAAGATTGCTACTGTTCTGTGATGACAATAC 1274
Db 434 GTCTGTGCTGGTGTCCATCTGACTTCAAAAAGATTGCTACTGTTCTGTGATGACAATAC 493
QY 1275 ACTAAATATCTGGCGCTTGAATAGAGGCTTAGAGAGAAACAGGAGGTGATAAACTTTC 1334
Db 494 ACTAAATATCTGGCGCTTGAATAGAGGCTTAGAGAGAAACAGGAGGTGATAAACTTTC 553
QY 1335 CACGGTGGGTTGGGCTCTCTCAGAGAAAAAGAGTCAAGACCTGGCTTAGTAAACAGTAAC 1394
Db 554 CACGGTGGGTTGGGCTCTCTCAGAGAAAAAGAGTCAAGACCTGGCTTAGTAAACAGTAAC 613
QY 1395 -GAGTAGCAGAGTACTCTCTGCAAAAGCCCCCA--GGGTAAAGTGCAATCCATCCAAATTC 1451
Db 614 TGAGTAGCAGAGTACTCTCTGCAAAAGCCCCCA--GGGTAAAGTGCAATCCATCCAAATTC 673
QY 1452 TTCCCCGTCATCCGAGCTTGTGCCCCAGCTGTGCT---GGAGACCTCCCTCTTCCCTT 1507
Db 674 TTCCCCGTCATCCGAGCTTGTGCCCCAGCTTGTGCCCCAGCTTGTGAGACCTCCCTCTTCCCTT 733
QY 1508 CAATACTCTCTACG--TTCTCTATTAAACCTCTCTCTGCAAGGCCGGTCTCCCATCAA 1565
Db 734 CAATACTCTCTACGTTTCTCTATTAAACCTCTCTCTGCAATGCCGGTCTCCCATCAA 792

RESULT 15
US-11-266-748A-322452/c
; Sequence 322452, Application US/11266748A
; Publication No. US20060134663A1
; GENERAL INFORMATION:
; APPLICANT: Harkin, Paul
; APPLICANT: Johnston, Patrick
; APPLICANT: Mulligan, Karl
; TITLE OF INVENTION: Transcriptome Microarray Technology and
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 55815-0102 (319189)
; CURRENT APPLICATION NUMBER: US/11/266,748A
; CURRENT FILING DATE: 2005-11-03
; PRIOR APPLICATION NUMBER: EP 04105479.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105482.6
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105483.4
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105507.0
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105485.9
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: EP 04105484.2
; PRIOR FILING DATE: 2004-11-03
; PRIOR APPLICATION NUMBER: US 60/662,276
```


;; PRIOR FILING DATE: 2005-03-14
;; PRIOR APPLICATION NUMBER: US 60/700,293
;; PRIOR FILING DATE: 2005-07-18
;; NUMBER OF SEQ ID NOS: 483996
;; SOFTWARE: PatentIn version 3.3
;; SEQ ID NO 322452
;; LENGTH: 798
;; TYPE: DNA
;; ORGANISM: Homo Sapiens
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (39)..(39)
;; OTHER INFORMATION: n is a, c, g, or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (71)..(71)
;; OTHER INFORMATION: n is a, c, g, or t
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (93)..(93)
;; OTHER INFORMATION: n is a, c, g, or t
US-11-266-748A-322452

Query Match 24.8%; Score 700.8; DB 8; Length 798;
Best Local Similarity 97.0%; Pred.No.2.5e-185;
Matches 756; Conservative 0; Mismatches 15; Indels 8; Gaps 4;

QY 795 TCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAGTATGGGA 854
DB 785 TCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAGTATGGGA 726

QY 855 TTTACGTAAGAATTATACCTTATCGACAAGAACCCATAGCATCCAAGTCTTTCTCTGTA 914
DB 725 TTTACGTAAGAATTATACCTTATCGACAAGAACCCATAGCATCCAAGTCTTTCTCTGTA 666

QY 915 CCCAGGTAGCAGCACTCGAAGAACTTTGGATATTCAGATCTGATTTGGATTCCACTGGCTC 974
DB 665 CCCAGGTAGCAGCACTCGAAGAACTTTGGATATTCAGATCTGATTTGGATTCCACTGGCTC 606

QY 975 TACTTTATTTGCTAAATTGACAGAGATAACATCTACATGTTTAAATATGACTGGGTTGAA 1034
DB 605 TACTTTATTTGCTAAATTGACAGAGATAACATCTACATGTTTAAATATGACTGGGTTGAA 546

QY 1035 GACTTCTCCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTATGTAAATCCAG 1094
DB 545 GACTTCTCCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTATGTAAATCCAG 486

QY 1095 CCTTAGTCCAGATGACCACTTTTATGTCAGTGGCTCAAGTGATGAAGCTGCCTCATATG 1154
DB 485 CCTTAGTCCAGATGACCACTTTTATGTCAGTGGCTCAAGTGATGAAGCTGCCTCATATG 426

QY 1155 GAAGGTCTCCACACCTGGCAACCTCTACTGTGCTCTGGGTCAATCTCAAGAGGTCAAC 1214
DB 425 GAAGGTCTCCACACCTGGCAACCTCTACTGTGCTCTGGGTCAATCTCAAGAGGTCAAC 366

QY 1215 GTCTGTGTGCTGGTGTCCATCTGACTTCAAAAGATTTGCTACTGTTCTGTATGACAAATAC 1274
DB 365 GTCTGTGTGCTGGTGTCCATCTGACTTCAAAAGATTTGCTACTGTTCTGTATGACAAATAC 306

QY 1275 ACTAAAAATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGAGGTGATAAACTTTC 1334
DB 305 ACTAAAAATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGAGGTGATAAACTTTC 246

QY 1335 CACGGTGGGTTGGGCTCTCAGAAGAAAAGAGTCAAGACCTGGCCTAGTAAACAGTAAC 1394
DB 245 CACGGTGGGTTGGGCTCTCAGAAGAAAAGAGTCAAGACCTGGCCTAGTAAACAGTAAC 186

QY 1395 -GAGTAGCCAGAGTACTCTGCCAAAGCCCCCA--GGGTAAAGTGCAATCCATCCAAATTC 1451
DB 185 TGAGTAGCCAGAGTACTCTGCCAAAGCCCCCAAGCGGTAAATGTGCAATCCATCCAAATTC 126

QY 1452 TTCCCCGTCATCCGAGCTTTGTGCCCAAGCTGTCT---GGAGACCTCCCTCTTCCTT 1507
|||||

Db 125 TTCCCCGTCATCCGAGCTTTGTGCCCAAGCTTTGGACGAGCTCCCTCTCTNTCCTT 66
QY 1508 CAAATACTCCTTACG-TTCTCTATTAAACCTTCTCTGCAAGGCCCGGTCTCCCATCAA 1565
|||||
Db 65 CAAATACTCCTTACGTTTCTCTATTAAACCTTCTCTGCAATGCCCGGTCTCCCATCAA 7
|||||

Search completed: November 9, 2006, 07:36:34
Job time : 645 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2006, 06:20:06 ; Search time 4564 Seconds

(without alignments)
7621.888 Million cell updates/sec

Title: US-10-726-160-1

Perfect score: 2831

Sequence: 1 ggcacgaggggttgag.....atgaaaaaaaaaaaaaa 2831

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 18892170 seqs, 6143817638 residues

Total number of hits satisfying chosen parameters: 37784340

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA Main:

- 1: /EMC_Celerra_SID33/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /EMC_Celerra_SID33/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 3: /EMC_Celerra_SID33/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 4: /EMC_Celerra_SID33/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 5: /EMC_Celerra_SID33/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 6: /EMC_Celerra_SID33/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 7: /EMC_Celerra_SID33/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 8: /EMC_Celerra_SID33/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 9: /EMC_Celerra_SID33/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 10: /EMC_Celerra_SID33/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 11: /EMC_Celerra_SID33/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 12: /EMC_Celerra_SID33/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 13: /EMC_Celerra_SID33/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
- 14: /EMC_Celerra_SID33/ptodata/2/pubpna/US11B_PUBCOMB.seq.*
- 15: /EMC_Celerra_SID33/ptodata/2/pubpna/US11C_PUBCOMB.seq.*
- 16: /EMC_Celerra_SID33/ptodata/2/pubpna/US11D_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2831	100.0	2831	7	US-10-409-511-1
2	2831	100.0	2831	9	US-10-726-160-1
3	2816.2	99.5	2838	3	US-09-780-053-1
4	2804	99.0	4422	3	US-09-814-353-20186
5	2804	99.0	4422	9	US-10-357-930-25699
6	2789.8	98.5	4221	6	US-10-145-396-28
7	2789.8	98.5	4221	7	US-10-393-590-85
8	2789.8	98.5	4221	7	US-10-393-567-85
9	2789.8	98.5	4221	7	US-10-394-087-85
10	2789.8	98.5	4221	7	US-10-172-118-1645
11	2789.8	98.5	4221	8	US-10-342-887-1645
12	2789.8	98.5	4221	10	US-10-848-755A-179
13	2442.4	86.3	4083	11	US-10-330-773-673
14	1520	53.7	1871	8	US-10-221-625-131
15	1302.2	46.0	2426	11	US-10-330-773-670
16	835.8	29.5	88892	11	US-10-330-773-672
17	649.8	23.0	57860	11	US-10-330-773-669

18	613	21.7	745	11	US-10-330-773-675	Sequence 675, App
19	558	19.7	708	6	US-10-076-555-715	Sequence 715, App
20	558	19.7	708	10	US-10-779-543-715	Sequence 715, App
21	507	17.9	507	6	US-10-206-901B-11	Sequence 11, Appl
22	500.4	17.7	815	10	US-10-779-543-3894	Sequence 3894, Ap
23	440	15.5	468	3	US-09-918-995-19957	Sequence 19957, A
24	439	15.5	543	3	US-09-814-353-16730	Sequence 16730, A
c	25	430.6	508	3	US-09-814-353-18083	Sequence 18083, A
26	421.4	14.9	445	3	US-09-780-053-3	Sequence 3, Appli
27	421.4	14.9	477	10	US-10-779-543-5959	Sequence 5959, Ap
c	28	411.6	434	3	US-09-814-353-5412	Sequence 5412, Ap
c	29	411.6	434	3	US-09-814-353-11699	Sequence 11699, A
30	403.4	14.5	493	6	US-10-106-698-491	Sequence 491, App
31	394.4	13.9	482	3	US-09-814-353-4038	Sequence 4038, Ap
32	394.4	13.9	482	3	US-09-814-353-10346	Sequence 10346, A
33	381.8	13.5	2001	8	US-10-403-571-73	Sequence 73, Appl
34	338	11.9	430	8	US-10-430-201-299	Sequence 299, App
35	338	11.9	430	8	US-10-430-201-300	Sequence 300, App
36	300	10.6	300	6	US-10-076-555-51	Sequence 51, Appl
37	300	10.6	300	10	US-10-779-543-51	Sequence 51, Appl
38	300	10.6	300	10	US-10-779-543-943	Sequence 943, App
39	300	10.6	300	10	US-10-779-543-1894	Sequence 1894, Ap
c	40	290.4	346	3	US-09-880-107-3634	Sequence 3634, Ap
41	282	10.0	297	9	US-10-357-930-8883	Sequence 8883, Ap
42	270	9.5	640	9	US-10-357-930-38800	Sequence 38800, A
43	254.4	9.0	432	16	US-11-128-061-1427	Sequence 1427, Ap
44	254.4	9.0	432	16	US-11-128-061-5069	Sequence 5069, Ap
45	254.4	9.0	432	16	US-11-128-049-1427	Sequence 1427, Ap

ALIGNMENTS

RESULT 1

US-10-409-511-1
; Sequence 1, Application US/10409511
; Publication No. US20030162268A1
; GENERAL INFORMATION:
; APPLICANT: IP, Nancy Y
; APPLICANT: CHEUNG, William M W
; TITLE OF INVENTION: No. US20030162268A1e1 Gene
; FILE REFERENCE: M98/0553/US
; CURRENT APPLICATION NUMBER: US/10/409,511
; CURRENT FILING DATE: 2003-04-08
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2831
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (124)..(2313)
US-10-409-511-1

Query Match	100.0%	Score 2831;	DB 7;	Length 2831;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2831;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	GGCACGAGCGGAGTTGGAGGCGATAACGATTGTGTGTGTGAGAGCGGCAACGTGCGATT	60	
Db	1	GGCACGAGCGGAGTTGGAGGCGATAACGATTGTGTGTGTGAGAGCGGCAACGTGCGATT	60	
QY	61	TCGTCTGAACCTGGAGGCAATTTCTACAGCTTTCTCTCAGCTGAGGCTTTCTCCGACC	120	
Db	61	TCGTCTGAACCTGGAGGCAATTTCTACAGCTTTCTCTCAGCTGAGGCTTTCTCCGACC	120	
QY	121	CTGATGCTCTTCAATTCCGTTGCTCCGACAGCCCGCTTGGCGCTCTGAGAAATGATGG	180	
Db	121	CTGATGCTCTTCAATTCCGTTGCTCCGACAGCCCGCTTGGCGCTCTGAGAAATGATGG	180	
QY	181	TCFTTCAATACCCCTCTTCAATCCCTTCTCAGCTGGTTATTCAGTGCAGTGGTAATGATGAA	240	
Db	181	TCFTTCAATACCCCTCTTCAATCCCTTCTCAGCTGGTTATTCAGTGCAGTGGTAATGATGAA	240	

Db 181 TCTTCAAAATACCCCTCTTCAATCCCTTCTGACTGGTTATCAGTGCAGTGGTAATGATGAA 240
Qy 241 CACACTTCTTATGAGAAACAGAGTCCAGAGTCCAGATTCTCTCTTTTGGATGATACCTTCTCTTCT 300
Db 241 CACACTTCTTATGAGAAACAGAGTCCAGAGTCCAGATTCTCTCTTTTGGATGATACCTTCTCTTCT 300
Qy 301 GCTCCCAATATGGAACATGATCTAGCAGTTCGCAATGGAAGAGGCTTCTTGGATGGTAT 360
Db 301 GCTCCCAATATGGAACATGATCTAGCAGTTCGCAATGGAAGAGGCTTCTTGGATGGTAT 360
Qy 361 AACACAGAAATCAAAAGTTTCAGAAAGAGTGTCTTCAAAAGAAATGGATGCTCACTGGAAAT 420
Db 361 AACACAGAAATCAAAAGTTTCAGAAAGAGTGTCTTCAAAAGAAATGGATGCTCACTGGAAAT 420
Qy 421 GCGCTCTTTGACCTGGCCCTGGTTCCTGGTGAACTTTAAATCTTGTTACAGCAGCAGGTGAT 480
Db 421 GCGCTCTTTGACCTGGCCCTGGTTCCTGGTGAACTTTAAATCTTGTTACAGCAGCAGGTGAT 480
Qy 481 CAAACAGCCAAATTTTGGGACGTAAAGCTGGTGAGCTGATTTGGAAATGCAAGGTCAAT 540
Db 481 CAAACAGCCAAATTTTGGGACGTAAAGCTGGTGAGCTGATTTGGAAATGCAAGGTCAAT 540
Qy 541 CAATGAGCTCAAGTCAGTTGCTTTTCTAAGTTTGAGAAAGCTGTATTTCTGTACGGGT 600
Db 541 CAATGAGCTCAAGTCAGTTGCTTTTCTAAGTTTGAGAAAGCTGTATTTCTGTACGGGT 600
Qy 601 GGAAGAGATGGCAACATTAATGTCTGGGATACAGGTGCAACAAAGAAATGGTTTTAT 660
Db 601 GGAAGAGATGGCAACATTAATGTCTGGGATACAGGTGCAACAAAGAAATGGTTTTAT 660
Qy 661 AGCAAGATGAATCAATCAGTGGAGCTCAATACCTCAGACAGCAAGCAACCCCTTCAAAA 720
Db 661 AGCAAGATGAATCAATCAGTGGAGCTCAATACCTCAGACAGCAAGCAACCCCTTCAAAA 720
Qy 721 CCCAAGAAAGAACAGAAATCAAAAGACTTGTCTCTTGTGGATTTCAGACAAAGTGT 780
Db 721 CCCAAGAAAGAACAGAAATCAAAAGACTTGTCTCTTGTGGATTTCAGACAAAGTGT 780
Qy 781 ACTGTGGTCTCTTTCAAGACGAGAAATCCTTAGTCTCAGCAGGAGCTGTGGATGGGATA 840
Db 781 ACTGTGGTCTCTTTCAAGACGAGAAATCCTTAGTCTCAGCAGGAGCTGTGGATGGGATA 840
Qy 841 ATCAAGATATGGGATTTACGTAGAATTTACTGCTTATCGACAGAAACCCATAGCATCC 900
Db 841 ATCAAGATATGGGATTTACGTAGAATTTACTGCTTATCGACAGAAACCCATAGCATCC 900
Qy 901 AGTCTTTTCTGTACCCAGGTAGCAGCTCGAAAACTTGGATTTCAAGTCTGATTTTG 960
Db 901 AGTCTTTTCTGTACCCAGGTAGCAGCTCGAAAACTTGGATTTCAAGTCTGATTTTG 960
Qy 961 GATTCACACTGGCTCTACTTTATTTGCTAATTTGCATATGACAGAGATAATCTACATGTTTAAAT 1020
Db 961 GATTCACACTGGCTCTACTTTATTTGCTAATTTGCATATGACAGAGATAATCTACATGTTTAAAT 1020
Qy 1021 ATGACTGGGTTGAAGACTTCTCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTT 1080
Db 1021 ATGACTGGGTTGAAGACTTCTCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTT 1080
Qy 1081 TATGTAATAATCCAGCTTAGTCCAGATGACAGTTTATTTAGTCAAGTGGCTCAAGTATGAA 1140
Db 1081 TATGTAATAATCCAGCTTAGTCCAGATGACAGTTTATTTAGTCAAGTGGCTCAAGTATGAA 1140
Qy 1141 GCTGCCCTACATATGGAAGTCTCCACACCTCGCAACCTCTACTGTGTCTCTGGGTCAAT 1200
Db 1141 GCTGCCCTACATATGGAAGTCTCCACACCTCGCAACCTCTACTGTGTCTCTGGGTCAAT 1200
Qy 1201 TCTCAAGAGTCAAGTCTGTGTCTGGTGTCTCCTGCTCAATCTGACTTCAAAAGATTTGCTCTGT 1260
Db 1201 TCTCAAGAGTCAAGTCTGTGTCTGGTGTCTCCTGCTCAATCTGACTTCAAAAGATTTGCTCTGT 1260
Qy 1261 TCTGATGACAAATACATTAATAATCTGGCGCTTGAATAGAGGCTTTAGAGGAGAAACCAAG 1320
Db 1261 TCTGATGACAAATACATTAATAATCTGGCGCTTGAATAGAGGCTTTAGAGGAGAAACCAAG 1320

Qy 1321 GGTGATAAATCTTCCACGGTGGGTTGGCCTCTCAGAGAGAAAAAGAGTCAAGACCTGGC 1380
Db 1321 GGTGATAAATCTTCCACGGTGGGTTGGCCTCTCAGAGAGAAAAAGAGTCAAGACCTGGC 1380
Qy 1381 CTAGTAAACAGTAAACAGAGTAGCCAGAGTACTCTCTGCAAAAGCCCCCAGGGTAAAGTGCAT 1440
Db 1381 CTAGTAAACAGTAAACAGAGTAGCCAGAGTACTCTCTGCAAAAGCCCCCAGGGTAAAGTGCAT 1440
Qy 1441 CCATCCAAATTTCTCCCCGTCAATCCGAGCTTGTGTGCCCCCAAGCTGTGTGGAGACCTCCCT 1500
Db 1441 CCATCCAAATTTCTCCCCGTCAATCCGAGCTTGTGTGCCCCCAAGCTGTGTGGAGACCTCCCT 1500
Qy 1501 CTTCTCTTCAATATCTTCAAGTCTCTATTAATAAAGCTCTCTGCAAGCCCGGTCTCCC 1560
Db 1501 CTTCTCTTCAATATCTTCAAGTCTCTATTAATAAAGCTCTCTGCAAGCCCGGTCTCCC 1560
Qy 1561 ATCAACAGAGAGGCTCTGTCTCCCTCTCTCCCAAGCCACCTTCAATCTTTCAAGATG 1620
Db 1561 ATCAACAGAGAGGCTCTGTCTCCCTCTCTCCCAAGCCACCTTCAATCTTTCAAGATG 1620
Qy 1621 TCGATTAGAAAATCTGGTGACCCGAAACACTTCTCTCATCAACCCCATCACTCCACTGCT 1680
Db 1621 TCGATTAGAAAATCTGGTGACCCGAAACACTTCTCTCATCAACCCCATCACTCCACTGCT 1680
Qy 1681 TCGGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCACTCC 1740
Db 1681 TCGGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCACTCC 1740
Qy 1741 CAAAGCAGAGGCTTGTCTGTAGTCTAGAAATAGAGTAAAGAGGAGCTAGACTCAAGCTGT 1800
Db 1741 CAAAGCAGAGGCTTGTCTGTAGTCTAGAAATAGAGTAAAGAGGAGCTAGACTCAAGCTGT 1800
Qy 1801 CTGGAGAGTGTGAAAACAAAAGTGTGAGAGTTGTAACTGTGTCACTGAGCTTGTATGGC 1860
Db 1801 CTGGAGAGTGTGAAAACAAAAGTGTGAGAGTTGTAACTGTGTCACTGAGCTTGTATGGC 1860
Qy 1861 CAAAGTTGAAAATCTTCAATTTGGATCTGTCTGCTTGTGTTAAACAGGAGACCTTAGT 1920
Db 1861 CAAAGTTGAAAATCTTCAATTTGGATCTGTCTGCTTGTGTTAAACAGGAGACCTTAGT 1920
Qy 1921 AAGGACTCTTAGTGTCTTACCAAAATCAAGCAAAAATGAAGAGAGCTGGTACCAAGTATCTCA 1980
Db 1921 AAGGACTCTTAGTGTCTTACCAAAATCAAGCAAAAATGAAGAGAGCTGGTACCAAGTATCTCA 1980
Qy 1981 GAGCCTCCGTCTCTPATCAGTCCGTATGCTTCAGAAAAGCTGTGGAAACGCTCTCTTCTCT 2040
Db 1981 GAGCCTCCGTCTCTPATCAGTCCGTATGCTTCAGAAAAGCTGTGGAAACGCTCTCTTCTCT 2040
Qy 2041 TTGAGACCTTGTGGAGAGGGTCTGAAATGGTAGCAAGAGATAGTTCCCCAGAGAT 2100
Db 2041 TTGAGACCTTGTGGAGAGGGTCTGAAATGGTAGCAAGAGATAGTTCCCCAGAGAT 2100
Qy 2101 AAAAACTGTTCTTGGCCATGGCAGCCAAAACGGAAGGCTGAGAACTCACTCTCCACGAAGT 2160
Db 2101 AAAAACTGTTCTTGGCCATGGCAGCCAAAACGGAAGGCTGAGAACTCACTCTCCACGAAGT 2160
Qy 2161 CCGTCATCCAGACACCCAAATTCAGGAGACAGAGCGGAAGACATTTGCAAGCCCGGTCT 2220
Db 2161 CCGTCATCCAGACACCCAAATTCAGGAGACAGAGCGGAAGACATTTGCAAGCCCGGTCT 2220
Qy 2221 ACCATCAGCCCGAGCTCCATGAGGAAATCTGCATATCTTCCATAGAAAGTCCCAAGAG 2280
Db 2221 ACCATCAGCCCGAGCTCCATGAGGAAATCTGCATATCTTCCATAGAAAGTCCCAAGAG 2280
Qy 2281 GACTTCTGTGGTCTGAAACACTCAACAGAAATATAGATTTCTAATCTGAGTGGTACTGA 2340
Db 2281 GACTTCTGTGGTCTGAAACACTCAACAGAAATATAGATTTCTAATCTGAGTGGTACTGA 2340
Qy 2341 GCTTTGGTCCACTAAAAAAGCTGAGCTTTGGTCCACTTAAACCAAGATGAAAAATACAAG 2400
Db 2341 GCTTTGGTCCACTAAAAAAGCTGAGCTTTGGTCCACTTAAACCAAGATGAAAAATACAAG 2400

Db 1261 TCTGATGACATACACTATAAAATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGGA 1320
QY 1321 GGTGATAAACTTTCCACGGTGGGTTGGGCTCTCAGAGAGAAAAAGAGTCAAGACCTGGC 1380
Db 1321 GGTGATAAACTTTCCACGGTGGGTTGGGCTCTCAGAGAGAAAAAGAGTCAAGACCTGGC 1380
QY 1381 CTAGTAAACAGTAAACAGTAGCCAGAGTACTCTGCGAAAGCCGCCAGGTAAGTGCAT 1440
Db 1381 CTAGTAAACAGTAAACAGTAGCCAGAGTACTCTGCGAAAGCCGCCAGGTAAGTGCAT 1440
QY 1441 CCATCAATTTCTCCCGCTCATCCGAGCTTGCGCCCAAGCTGTGCGAGACCTCCCT 1500
Db 1441 CCATCAATTTCTCCCGCTCATCCGAGCTTGCGCCCAAGCTGTGCGAGACCTCCCT 1500
QY 1501 CTTCTTCAAAATACCTCTAGTTCTCTATTAAAACTCTCTGCCAAGCCCGGCTCC 1560
Db 1501 CTTCTTCAAAATACCTCTAGTTCTCTATTAAAACTCTCTGCCAAGCCCGGCTCC 1560
QY 1561 ATCAACAGAGAGGCTCTGCTCTCCCGTCTCTCCAGCCACCTTCACTTTCAAGATG 1620
Db 1561 ATCAACAGAGAGGCTCTGCTCTCCCGTCTCTCCAGCCACCTTCACTTTCAAGATG 1620
QY 1621 TCGATTAGAAACTGGGTGACCCGACACCTTCTCATCACACCCATCACTCCACTGCT 1680
Db 1621 TCGATTAGAAACTGGGTGACCCGACACCTTCTCATCACACCCATCACTCCACTGCT 1680
QY 1681 TCGGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTCTGTGAGCCAGAGTCACTCC 1740
Db 1681 TCGGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTCTGTGAGCCAGAGTCACTCC 1740
QY 1741 CAAGCAGAGGCTGCTCTGAGTCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGT 1800
Db 1741 CAAGCAGAGGCTGCTCTGAGTCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGT 1800
QY 1801 CTGGAGAGTGTAACAAAGTGTTGAAGAGTTGTAACGTGTGACTGAGCTTCATGTC 1860
Db 1801 CTGGAGAGTGTAACAAAGTGTTGAAGAGTTGTAACGTGTGACTGAGCTTCATGTC 1860
QY 1861 CAAGTTGAAAATCTTTCATTGGATCTGTGCTGCTGCTTGGTGAACAGGAGACCTTAGT 1920
Db 1861 CAAGTTGAAAATCTTTCATTGGATCTGTGCTGCTGCTTGGTGAACAGGAGACCTTAGT 1920
QY 1921 AAGGACTCTTAGTCTTACCAAAATCAAGCAAAATTAAGAGAGGCTGGTACCAAGTATCTCA 1980
Db 1921 AAGGACTCTTAGTCTTACCAAAATCAAGCAAAATTAAGAGAGGCTGGTACCAAGTATCTCA 1980
QY 1981 GAGCTCTCGTCTCTATCAGTCCGTATGCTTCAGAAAAGCTGTGGAACGCTACTCTTCT 2040
Db 1981 GAGCTCTCGTCTCTATCAGTCCGTATGCTTCAGAAAAGCTGTGGAACGCTACTCTTCT 2040
QY 2041 TTGAGACCTTTGGAGAGGGTCTGAAATGGTAGGCAAAAGAGAAATAGTTTCCCGAGAGT 2100
Db 2041 TTGAGACCTTTGGAGAGGGTCTGAAATGGTAGGCAAAAGAGAAATAGTTTCCCGAGAGT 2100
QY 2101 AAAAACTGGTTGTGGCCATGGCAGCAAAACGAGGCTGAGAAATCCATCTCCACGAAGT 2160
Db 2101 AAAAACTGGTTGTGGCCATGGCAGCAAAACGAGGCTGAGAAATCCATCTCCACGAAGT 2160
QY 2161 CCGTCAATCCAGACACCAATTCAGAGGACAGAGCGGAAAGACATTCGCAAGCCGGTC 2220
Db 2161 CCGTCAATCCAGACACCAATTCAGAGGACAGAGCGGAAAGACATTCGCAAGCCGGTC 2220
QY 2221 ACCATCAGCCCGAGCTCCATGAGGAAAATCTGCACATACTTCCATAGAAAAGTCCAGGAG 2280
Db 2221 ACCATCAGCCCGAGCTCCATGAGGAAAATCTGCACATACTTCCATAGAAAAGTCCAGGAG 2280
QY 2281 GACTTCTGTGCTTGAACACTCAACAGAAATATAGATTTCTATCTGAGTGAAGTACTGA 2340
Db 2281 GACTTCTGTGCTTGAACACTCAACAGAAATATAGATTTCTATCTGAGTGAAGTACTGA 2340
QY 2341 GCTTTGGTCCACTAAAAACAGCTTGAGCTTTGGTCCACTAAATCAAGATGAAAAATACAG 2400

Db 2341 GCTTTGGTCCACTAAAAACAGCTTGAGCTTTGGTCCACTAAAAACAGATGAAAAATACAG 2400
QY 2401 AGTGACTCTATAACTCTGGTCTTTAAGAAAGCTGCTTTTCATTTTATAGACAAATCTTT 2460
Db 2401 AGTGACTCTATAACTCTGGTCTTTAAGAAAGCTGCTTTTCATTTTATAGACAAATCTTT 2460
QY 2461 TCAACGCTGAAATGTACCTAATCTGGTCTACTACTACCAATAATGTATATGCAGCTTCCCAG 2520
Db 2461 TCAACGCTGAAATGTACCTAATCTGGTCTACTACTACCAATAATGTATATGCAGCTTCCCAG 2520
QY 2521 GATGAATGCTGCTGTTTAAATTTCAAAAGTAAATTTGTCACCTAGCAATTTGAATGAAT 2580
Db 2521 GATGAATGCTGCTGTTTAAATTTCAAAAGTAAATTTGTCACCTAGCAATTTGAATGAAT 2580
QY 2581 AGTCTTCACTTTTAAATTTTCAATCTCTCTATAATAATGACATCCAGTTCATGGAGG 2640
Db 2581 AGTCTTCACTTTTAAATTTTCAATCTCTCTATAATAATGACATCCAGTTCATGGAGG 2640
QY 2641 CAAAAACAAAGTTTCTGTTATCTGTAACCTTCTATGCTCAGTGGAAAGTATCTGCCAG 2700
Db 2641 CAAAAACAAAGTTTCTGTTATCTGTAACCTTCTATGCTCAGTGGAAAGTATCTGCCAG 2700
QY 2701 CCACAGCATGAGGCTGTGAAAGGCTGACTGAGAAATCCTCTGCTGAAGACCCCTGGTTCT 2760
Db 2701 CCACAGCATGAGGCTGTGAAAGGCTGACTGAGAAATCCTCTGCTGAAGACCCCTGGTTCT 2760
QY 2761 GTTCTGCCCTCAACATGTATTAATTTTATTGAAATACATAATCTTTTCACTATGAAAAA 2820
Db 2761 GTTCTGCCCTCAACATGTATTAATTTTATTGAAATACATAATCTTTTCACTATGAAAAA 2820
QY 2821 AAAAAAATAA 2831
Db 2821 AAAAAAATAA 2831

RESULT 3
US-09-780-053-1
; Sequence 1, Application US/097800053
; Patent No. US20020102640A1
; GENERAL INFORMATION:
; APPLICANT: Rene S. Hubert
; APPLICANT: Daniel E.H. Afar
; APPLICANT: Pia M. Challita-Bid
; APPLICANT: Mary Paris
; APPLICANT: Elana Levin
; APPLICANT: Steve Chappell Mitchell
; APPLICANT: Aya Jakobovits
; TITLE OF INVENTION: 83P5G4: A TISSUE SPECIFIC PROTEIN
; FILE REFERENCE: 129,5USU1
; CURRENT APPLICATION NUMBER: US/09/780,053
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/181,261
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 716
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2838
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (130)...(2322)
US-09-780-053-1

Query Match 99.5%; Score 2816.2; DB 3; Length 2838;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2821; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 3 CACGAGCGGGAGTTGGAGCGCAATAACGATTTTGTGTGTGAGAGCGCAACGTGCCATTC 62
|||
Db 9 CAGTGGCGGAGTTGGAGCGCAATAACGATTTTGTGTGTGAGAGCGCAACGTGCCATTC 68
|||

QY 63 TGCTGAACCTGGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCT 122
DB 69 TGCTGAACCTGGAGGCAATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCT 128
QY 123 GATGCTCTTCAATTCGGTGCTCGCGCAGGCCCAAGCTTGGCGTCTCAGAAAATGGATGGTC 182
DB 129 GATGCTCTTCAATTCGGTGCTCGCGCAGGCCCAAGCTTGGCGTCTCAGAAAATGGATGGTC 188
QY 183 TTCCAAATACCCCTTCTCAATCCCTCTGACTGGTATCAGTCAGAGGGTAAATGATGAACA 242
DB 189 TTCCAAATACCCCTTCTCAATCCCTCTGACTGGTATCAGTCAGAGGGTAAATGATGAACA 248
QY 243 CACTTCTTATGGAGAAACAGGAGTCCAGTCTCCTCTTTTGGATGTACCTTCTCTCTCGC 302
DB 249 CACTTCTTATGGAGAAACAGGAGTCCAGTCTCCTCTTTTGGATGTACCTTCTCTCTCGC 308
QY 303 TCCCAATATGGAACATGTACTAGCAGTTGCGCAATGAAGAAGCTTTGGTTCGATTTGATAA 362
DB 309 TCCCAATATGGAACATGTACTAGCAGTTGCGCAATGAAGAAGCTTTGGTTCGATTTGATAA 368
QY 363 CACAGAAATCAAAAGTTTCAGAAAGAGTGCCTTCAAGAAATGGATGGCTCACTGGAATGC 422
DB 369 CACAGAAATCAAAAGTTTCAGAAAGAGTGCCTTCAAGAAATGGATGGCTCACTGGAATGC 428
QY 423 CGTCTTTGACCTGGCCTGGGTTCTCGGTGAACCTTAAACTTTGTTACAGCAGCAGGTGATCA 482
DB 429 CGTCTTTGACCTGGCCTGGGTTCTCGGTGAACCTTAAACTTTGTTACAGCAGCAGGTGATCA 488
QY 483 AACAGCCAAATTTTGGGAGCTAAAAGCTGGTGAAGTGTATGGAAACATGCAAAAGGTCA 542
DB 489 AACAGCCAAATTTTGGGAGCTAAAAGCTGGTGAAGTGTATGGAAACATGCAAAAGGTCA 548
QY 543 ATGCAGCCTCAAGTCAGTGGCTTTTCTTAAGTTTGAGAAAGCTGTATTTCTGACGGGTGG 602
DB 549 ATGCAGCCTCAAGTCAGTGGCTTTTCTTAAGTTTGAGAAAGCTGTATTTCTGACGGGTGG 608
QY 603 AAGAGATGGCAACATTTATGTCTGGGATACACAGTGCACAAAGAAAGATGGGTTTTATAG 662
DB 609 AAGAGATGGCAACATTTATGTCTGGGATACACAGTGCACAAAGAAAGATGGGTTTTATAG 668
QY 663 GCAAGTGAATCAATCAGTGGAGCTCACAATACCTCAGCAAGCAAAACCCCTTCAAAACC 722
DB 669 GCAAGTGAATCAATCAGTGGAGCTCACAATACCTCAGCAAGCAAAACCCCTTCAAAACC 728
QY 723 CAAGAAGAAACAGAAATCAAAAGGACTTGTCTCTCTGTGGATTTCCAGCAAGAGTGTAC 782
DB 729 CAAGAAGAAACAGAAATCAAAAGGACTTGTCTCTCTGTGGATTTCCAGCAAGAGTGTAC 788
QY 783 TGTGGTCTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAAT 842
DB 789 TGTGGTCTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAAT 848
QY 843 CAAAGTATGGGATTTACGTAAAGAAATATATCTGCTTATCGCAAGAACCCATAGCATCCAA 902
DB 849 CAAAGTATGGGATTTACGTAAAGAAATATATCTGCTTATCGCAAGAACCCATAGCATCCAA 908
QY 903 GTCTTTCTGTACCCAGGTAGCAGCATCGAAAACCTTGGATATTCAGTCTGATTTTGA 962
DB 909 GTCTTTCTGTACCCAGGTAGCAGCATCGAAAACCTTGGATATTCAGTCTGATTTTGA 968
QY 963 TTCCACTGGCTCTACCTTTATTTGCTAAATTTGCAAGAGTAAATCTACATCTGATTTTAAAT 1022
DB 969 TTCCACTGGCTCTACCTTTATTTGCTAAATTTGCAAGAGTAAATCTACATGTTTAAATAT 1028
QY 1023 GACTGGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACCAAGAACTCTACCTTTTA 1082
DB 1029 GACTGGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACCAAGAACTCTACCTTTTA 1088
QY 1083 TGTAAATCCAGCCTTAGTCCAGATGACCAAGTTTTTATGTCAGTGGCTCAAGTATGAAGC 1142
DB 1089 TGTAAATCCAGCCTTAGTCCAGATGACCAAGTTTTTATGTCAGTGGCTCAAGTATGAAGC 1148
QY 1143 TGCTACATATGGAAGGTCTCCACACCCCTGGCAACCTCTCTACTGTGCTCTCTGGTCAATTC 1202

DB 1149 TGCTACATATGGAAGGTCTCCACACCCCTGGCAACCTCTCTACTGTGCTCTCTGGTCAATTC 1208
QY 1203 TCAAGAGGTCAAGTCTGTGTGCTGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1262
DB 1209 TCAAGAGGTCAAGTCTGTGTGCTGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1268
QY 1263 TGATCACAATACACATTAATAAATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGGAGG 1322
DB 1269 TGATCACAATACACATTAATAAATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGGAGG 1328
QY 1323 TGATAAACTTTTCCACGGTGGGTGGGCTCTCAGAAAGAAAAAGAGTCAAGACCTTGGGCT 1382
DB 1329 TGATAAACTTTTCCACGGTGGGTGGGCTCTCAGAAAGAAAAAGAGTCAAGACCTTGGGCT 1388
QY 1383 AGTAACAGTACAGTACAGTACAGTACCTCTCTGCAAGCCCGCCAGGGTAAAGTGAATCC 1442
DB 1389 AGTAACAGTACAGTACAGTACCTCTCTGCAAGCCCGCCAGGGTAAAGTGAATCC 1448
QY 1443 ATCCAAATTTCTCCCGTCAATCCGAGCTTGTGCCCCCAAGCTGTGTGAGAGCTCTCCCTCT 1502
DB 1449 ATCCAAATTTCTCCCGTCAATCCGAGCTTGTGCCCCCAAGCTGTGTGAGAGCTCTCCCTCT 1508
QY 1503 TCCCTTCAAATACCTCTACGTTCTTATTAAAACTCTCTCTGCAAGCCCGGTCTCCCAT 1562
DB 1509 TCCCTTCAAATACCTCTACGTTCTTATTAAAACTCTCTCTGCAAGCCCGGTCTCCCAT 1568
QY 1563 CAAACAGAGAGGCTGTCT 1622
DB 1569 CAAACAGAGAGGCTGTCT 1628
QY 1623 GATTAGAACTGGGTGACCCGAAACCTTCTCTCATCACCCCATCACCTCCACCTGCTTC 1682
DB 1629 GATTAGAACTGGGTGACCCGAAACCTTCTCTCATCACCCCATCACCTCCACCTGCTTC 1688
QY 1683 GGAGACCAAGATCATGTCTCCGAGAAAAACCCCTTATCTCTGTGAGCCAGAACTCATCCCA 1742
DB 1689 GGAGACCAAGATCATGTCTCCGAGAAAAACCCCTTATCTCTGTGAGCCAGAACTCATCCCA 1748
QY 1743 AGCAGAGGCTTGTCTGTAGTCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCT 1802
DB 1749 AGCAGAGGCTTGTCTGTAGTCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCT 1808
QY 1803 GGAGAGTGTGAACAAAAAGTGTGCAAGAGTTGTAACCTGTGTGACTGTAGCTTGTATGGCCA 1862
DB 1809 GGAGAGTGTGAACAAAAAGTGTGCAAGAGTTGTAACCTGTGTGACTGTAGCTTGTATGGCCA 1868
QY 1863 AGTTGAAAAATCTTCATTTGGATCTGTGCTGCTCTCTGTGTAAACCCAGGAAGACCTTAGTAA 1922
DB 1869 AGTTGAAAAATCTTCATTTGGATCTGTGCTGCTCTGTGTAAACCCAGGAAGACCTTAGTAA 1928
QY 1923 GGACTCTCTAGGCTCCTACCAATCAAGCAAAATGGAAGAGCTGGTACAGATATCTCAGA 1982
DB 1929 GGACTCTCTAGGCTCCTACCAATCAAGCAAAATGGAAGAGCTGGTACAGATATCTCAGA 1988
QY 1983 GCCTCCGCTCTCTATCAGTCCGCTATGCTTCAAGAAAGCTGTGAAACGCTACCTCTCTCTTT 2042
DB 1989 GCCTCCGCTCTCTATCAGTCCGCTATGCTTCAAGAAAGCTGTGAAACGCTACCTCTCTCTTT 2048
QY 2043 GAGACTCTGTGAGAAAGGCTCTGAAATGGTAGGCAAGAAAGATAGTTTCCCCAGAGAAATAA 2102
DB 2049 GAGACTCTGTGAGAAAGGCTCTGAAATGGTAGGCAAGAAAGATAGTTTCCCCAGAGAAATAA 2108
QY 2103 AAACTGGTGTGGGCAATGGGCAAGCAAAACGGAAGGCTGAGAAATCCATCTCCACGAAGTCC 2162
DB 2109 AAACTGGTGTGGGCAATGGGCAAGCAAAACGGAAGGCTGAGAAATCCATCTCCACGAAGTCC 2168
QY 2163 GTCATCCAGACACCAATTTCCAGAGCAGAGCGGAAGACATTTGCCAAGCCCGGTAC 2222
DB 2169 GTCATCCAGACACCAATTTCCAGAGCAGAGCGGAAGACATTTGCCAAGCCCGGTAC 2228
QY 2223 CATCACGCCAGCTCCATGAGGAAATCTGCACATATTTCCATAGAAAAGTCCAGGAGGA 2282

Db 2229 CATCAGCCCGAGCTCCATGAGGAAATCTGCACATACCTTCCATAGAAAGTCCCGAGGGA 2288
QY 2283 CTTCTGTGTCTGTAACACTCAACAGAAATATAGATTTCTAAATCTGAGTGAGTTACTGAGC 2342
Db 2289 CTTCTGTGTCTGTAACACTCAACAGAAATATAGATTTCTAAATCTGAGTGAGTTACTGAGC 2348
QY 2343 TTTTGTGCCATAAACAAGCTGAGCTTGGTCCACTAAACAGATGAAATATACAGAG 2402
Db 2349 TTTTGTGCCATAAACAAGCTGAGCTTGGTCCACTAAACAGATGAAATATACAGAG 2408
QY 2403 TGACTCTATAACTCTGTGCTTTTAAAGAAAGCTGCTTTTCAATTTTATAGACAAATCTTTTC 2462
Db 2409 TGACTCTATAACTCTGTGCTTTTAAAGAAAGCTGCTTTTCAATTTTATAGACAAATCTTTTC 2468
QY 2463 AACGCTGAAATGATCAATCTGGTTCCTACTACCAATAATGTATATGACAGCTTCCCGAGGA 2522
Db 2469 AACGCTGAAATGATCAATCTGGTTCCTACTACCAATAATGTATATGACAGCTTCCCGAGGA 2528
QY 2523 TGAATGCTGTGTTTAAATTTCAATGTAATTTTGTCACTCTAGCATTTTGAATGAATAG 2582
Db 2529 TGAATGCTGTGTTTAAATTTCAATGTAATTTTGTCACTCTAGCATTTTGAATGAATAG 2588
QY 2583 TCTTCACTTTTAAATTTTCACTTCTCTATAATAATGACATCCAGTTTCAATGAGGCA 2642
Db 2589 TCTTCACTTTTAAATTTTCACTTCTCTATAATAATGACATCCAGTTTCAATGAGGCA 2648
QY 2643 AAAAAACAAGTTCTTGTGTTATCTCGAAACTTTTCTATGCTCAGTGGAAAGTATCTGCCAGCC 2702
Db 2649 AAAAAACAAGTTCTTGTGTTATCTCGAAACTTTTCTATGCTCAGTGGAAAGTATCTGCCAGCC 2708
QY 2703 ACAGCATGAGGCTGTGAAGGCTGACTGAGAAATCTCTGCTGAAGACCCCTGTTCTGT 2762
Db 2709 ACAGCATGAGGCTGTGAAGGCTGACTGAGAAATCTCTGCTGAAGACCCCTGTTCTGT 2768
QY 2763 TCTGCTCCCAACATGTATATTTTATTTGAAATACATATCTTTTCACTATGAAAAAAA 2822
Db 2769 TCTGCTCCCAACATGTATATTTTATTTGAAATACATATCTTTTCACTATGAAAAAAA 2828
QY 2823 AAAAAAAA 2831
Db 2829 AAAAAAAA 2837

RESULT 4
US-09-814-353-20186
; Sequence 20186, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20186
; LENGTH: 4422

; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: 4406, 4407, 4408, 4409, 4410, 4411, 4412, 4413, 4414, 4415,
; LOCATION: 4416, 4417, 4418, 4419, 4420, 4421, 4422
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-20186
Query Match 99.0%; Score 2804; DB 3; Length 4422;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2807; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 3 CACGAGCGGAGTTGGAGCGCATACGATTTGTGTGTGAGAGGGCGAACCTGGGATTTTC 62
Db 8 CAGTGGCGGGAGTTGGAGCGCATACGATTTGTGTGTGAGAGGGCGAACCTGGGATTTTC 67
QY 63 TGCTGAACCTCGAGGCACTTTCTAGACCTTTCTCTCAGCTCAGGCTTTTCTCCGACCT 122
Db 68 TGCTGAACCTCGAGGCACTTTCTAGACCTTTCTCTCAGCTCAGGCTTTTCTCCGACCT 127
QY 123 GATGCTCTTCAATTCGGTGTCTCCGACCCAGCCCGCTTGGCGTCTGAGAAATGGAATGTC 182
Db 128 GATGCTCTTCAATTCGGTGTCTCCGACCCAGCCCGCTTGGCGTCTGAGAAATGGAATGTC 187
QY 183 TTCAATACCTCTTCAATCTCTGACTGGTTATCAGTGCGAGTGGTAAATGATGAACA 242
Db 188 TTCAATACCTCTTCAATCTCTGACTGGTTATCAGTGCGAGTGGTAAATGATGAACA 247
QY 243 CACTTCTTATGAGAAACAGGAGTCCAGTTTCTCTCTTTGGATGTACCTTCTCTCTGC 302
Db 248 CACTTCTTATGAGAAACAGGAGTCCAGTTTCTCTCTTTGGATGTACCTTCTCTCTGC 307
QY 303 TCCCAATATGGAACATGTACTAGCAGTTGCCAATGAAGAGCTTTGTTCGATTTGTATAA 362
Db 308 TCCCAATATGGAACATGTACTAGCAGTTGCCAATGAAGAGCTTTGTTCGATTTGTATAA 367
QY 363 CACAGAAATCAAAAGTTTCAGAAAGAGTGTCTCAAGAAATGGAATGGTCTCACTGGAATGC 422
Db 368 CACAGAAATCAAAAGTTTCAGAAAGAGTGTCTCAAGAAATGGAATGGTCTCACTGGAATGC 427
QY 423 CGTCTTTGACCTGGCTGGGTTCTGTTGAACTTTAAACTTTTACAGCAGCAGGTGATCA 482
Db 428 CGTCTTTGACCTGGCTGGGTTCTGTTGAACTTTAAACTTTTACAGCAGCAGGTGATCA 487
QY 483 AACAGCCAAATTTTGGGACGTAAAGCTGGTGAATTTGGAAATGGAAATGCAAGTCAATCA 542
Db 488 AACAGCCAAATTTTGGGACGTAAAGCTGGTGAATTTGGAAATGCAAGTCAATCA 547
QY 543 ATGAGCCTCAAGTCAGTTGCTTTTCTAAGTTTGAAGAAAGCTGATTTCTGACGGTGG 602
Db 548 ATGAGCCTCAAGTCAGTTGCTTTTCTAAGTTTGAAGAAAGCTGATTTCTGACGGTGG 607
QY 603 AAGAGATGGCAACATTATGTTCTGGATACCAAGTGGCAACAAAAGATGGGTTTATAG 662
Db 608 AAGAGATGGCAACATTATGTTCTGGATACCAAGTGGCAACAAAAGATGGGTTTATAG 667
QY 663 GCAAGTGAATCAAAATCAGTGGAGCTCACAATACCTCAGCAAGCAAAACCCCTTCAAAACC 722
Db 668 GCAAGTGAATCAAAATCAGTGGAGCTCACAATACCTCAGCAAGCAAAACCCCTTCAAAACC 727
QY 723 CAAGAAAGAAACAGAAATTCAAAAGGCTTCTCTGTTGAAATTTCCAGCAAGGATTAAT 782
Db 728 CAAGAAAGAAACAGAAATTCAAAAGGCTTCTCTGTTGAAATTTCCAGCAAGGATTAAT 787
QY 783 TGTGCTCTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGATTAAT 842
Db 788 TGTGCTCTCTTTCAAGACGAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGATTAAT 847
QY 843 CAAAGTATGGGATTTACGTAAGAAATTTATCTGCTTATCGACAAAGCAACCCATAGCATCCAA 902
Db 848 CAAAGTATGGGATTTACGTAAGAAATTTATCTGCTTATCGACAAAGCAACCCATAGCATCCAA 907

Qy	903	GTCTTTCTCTGTA	CCGAGGTAGCAGCACTCGAAAACTTGGGATATTTCAAGTCTGATTTTGGG	962
Db	908	GTCTTTCTCTGTAC	CCAGGTAGCAGCACTCGAAAACTTGGATATTTCAAGTCTGATTTTGGG	967
Qy	963	TTCCACTGGGCTCTACT	TTTATTTTGTCTAAATTGCACAGACGATAAATCTACATGATGTTTAATAT	1022
Db	968	TTCCACTGGGCTCTACT	TTTATTTTGTCTAAATTGCACAGACGATAAATCTACATGATGTTTAATAT	1027
Qy	1023	GACTGGGTTGAAGACTT	CTCCAGTGGCTATTTTCAATGGACACGAACTCTACCTTTTAA	1082
Db	1028	GACTGGGTTGAAGACTT	CTCCAGTGGCTATTTTCAATGGACACGAACTCTACCTTTTAA	1087
Qy	1083	TGTAAATCCAGCCCTTAG	TCAGATGACAGTCTTTTATGTCAGTGGCTCAAGTGAATGAAGC	1142
Db	1088	TGTAAATCCAGCCCTTAG	TCAGATGACAGTCTTTTATGTCAGTGGCTCAAGTGAATGAAGC	1147
Qy	1143	TGCTCATATGGAAGGCT	CTCCACACCTGGCAACCTCTCTACCTTTTCTCTGGGTCAATC	1202
Db	1148	TGCTCATATGGAAGGCT	CTCCACACCTGGCAACCTCTCTACCTTTTCTCTGGGTCAATC	1207
Qy	1203	TCAGAGGTCACGTCCTG	TGTGCTGGTGCATCTGCACTTCAAAAGATTTGCTACTCTGTC	1262
Db	1208	TCAGAGGTCACGTCCTG	TGTGCTGGTGCATCTGCACTTCAAAAGATTTGCTACTCTGTC	1267
Qy	1263	TGATGACAATACACTAA	AAAAATCTGGCGCTTGAATAGAGGCTTAGAGAGAGAAACGAGAGG	1322
Db	1268	TGATGACAATACACTAA	AAAAATCTGGCGCTTGAATAGAGGCTTAGAGAGAGAAACGAGAGG	1327
Qy	1323	TGATAAACTTTCCACG	TGGGTTGGGCTCTCAGAGAGAAAGAGTCAGACCTGGGCT	1382
Db	1328	TGATAAACTTTCCACG	TGGGTTGGGCTCTCAGAGAGAAAGAGTCAGACCTGGGCT	1387
Qy	1383	AGTAAACAGTAAACAG	TAGCAGAGTACTCTGCCAAAGCCGCCAGGGTAAAGTGCATATCC	1442
Db	1388	AGTAAACAGTAAACAG	TAGCAGAGTACTCTGCCAAAGCCGCCAGGGTAAAGTGCATATCC	1447
Qy	1443	ATCCAAATCTTCCCG	TCATCCGCGTGTGCCCCAAGCTGTGCTGGAGACCTTCCCTCT	1502
Db	1448	ATCCAAATCTTCCCG	TCATCCGCGTGTGCCCCAAGCTGTGCTGGAGACCTTCCCTCT	1507
Qy	1503	TCCTTCAAAATCTCTA	CGTTCTCTATAAACCCTCTCTGCCAAGCCCGGTCTCCCAT	1562
Db	1508	TCCTTCAAAATCTCTA	CGTTCTCTATAAACCCTCTCTGCCAAGCCCGGTCTCCCAT	1567
Qy	1563	CAACAGAAAGAGGCTC	TGCTCTCCGTCCTCTCCCAAGCCACTTTCATCTTTCAAGATGTC	1622
Db	1568	CAACAGAAAGAGGCTC	TGCTCTCCGTCCTCTCCCAAGCCACTTTCATCTTTCAAGATGTC	1627
Qy	1623	GATTAGAAACTGGGTG	ACCCGAAACACCTTCTCTATCACCCCACTCATCCACCTGCTTC	1682
Db	1628	GATTAGAAACTGGGTG	ACCCGAAACACCTTCTCTATCACCCCACTCATCCACCTGCTTC	1687
Qy	1683	GGAGACCAAGATCATG	TCTCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCAATCCCA	1742
Db	1688	GGAGACCAAGATCATG	TCTCGAGAAAGCCCTTATTTCTGTGAGCCAGAGTCAATCCCA	1747
Qy	1743	AGCAGAGGCTTGCTCT	CGAGTCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCT	1802
Db	1748	AGCAGAGGCTTGCTCT	CGAGTCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCT	1807
Qy	1803	GGAGAGTGTGAAACAA	AAAGTGTGGAAGAGTGTAACTGTGTGACTGAGCTTGAATGGCCA	1862
Db	1808	GGAGAGTGTGAAACAA	AAAGTGTGGAAGAGTGTAACTGTGTGACTGAGCTTGAATGGCCA	1867
Qy	1863	AGTTGAAAATCTTCAT	TTTGGATCTGTGCTGCTTGTGGTAAACAGAGAGACCTTAGTAA	1922
Db	1868	AGTTGAAAATCTTCAT	TTTGGATCTGTGCTGCTTGTGGTAAACAGAGAGACCTTAGTAA	1927
Qy	1923	GGACTCTCTAGGTCCT	ACCACAAATCAAGCAAAATTTGAAGGAGCTGGTACCAGTATCTCAGA	1982
Db	1928	GGACTCTCTAGGTCCT	ACCACAAATCAAGCAAAATTTGAAGGAGCTGGTACCAGTATCTCAGA	1987
Qy	1983	GCCTCCGCTCTCCTAT	CAAGTCCGATGCTTTCAGAAAGCTGTGGAAACGCTACTCTTCCCTTT	2042

1988	Db		GCCTCCGTCCTTATCAGTCCGATAGCTTCAGAAAGCTGTGGAAACGCTACTCTTCTCTTT	2047
2043	Qy	GAGACCTTGTGGAGAGGGTCTGAAATGGTAGGCAAGAGAAATAGTTCCTCCAGAGATAA	2102	
2048	Db	GAGACCTTGTGGAGAGGGTCTGAAATGGTAGGCAAGAGAAATAGTTCCTCCAGAGATAA	2107	
2103	Qy	AAACTGGTTGTGGCCATGGCAGCCAAACGGAAGGCTGAGAATCCATCTCCACGAAGTCC	2162	
2108	Db	AAACTGGTTGTGGCCATGGCAGCCAAACGGAAGGCTGAGAATCCATCTCCACGAAGTCC	2167	
2163	Qy	GTCAATCCAGACACCCAAATTCAGAGGACACAGAGCGGGAAGACATTCGCCAGCCCGGTCA	2222	
2168	Db	GTCAATCCAGACACCCAAATTCAGAGGACACAGAGCGGGAAGACATTCGCCAGCCCGGTCA	2227	
2223	Qy	CATCAGCCCCAGCTCCATGAGGAAATCTGCACATCTCTCCATAGAAGTCCCGAGGAGGA	2282	
2228	Db	CATCAGCCCCAGCTCCATGAGGAAATCTGCACATCTCTCCATAGAAGTCCCGAGGAGGA	2287	
2283	Qy	CTTCTGTGGTCTCGAACACCTCAACAGAAATATAGATTTCTAATCTGAGTGAAGTTACTGAGC	2342	
2288	Db	CTTCTGTGGTCTCGAACACCTCAACAGAAATATAGATTTCTAATCTGAGTGAAGTTACTGAGC	2347	
2343	Qy	TTTGGTCCACTAAAAAAGCTGAGCTTTGGTCCACTTAAAAACAAGATGAAAAATACAAGAG	2402	
2348	Db	TTTGGTCCACTAAAAAAGCTGAGCTTTGGTCCACTTAAAAACAAGATGAAAAATACAAGAG	2407	
2403	Qy	TGACTCTATAACTCTGGTCTTTAAGAAAGCTGCTTTTCAATTTTACAGAAAATCTTTTC	2462	
2408	Db	TGACTCTATAACTCTGGTCTTTAAGAAAGCTGCTTTTCAATTTTACAGAAAATCTTTTC	2467	
2463	Qy	AACGCTGAAATGTACCTAATCTGGTCTCTACTACCAATAGTATGACAGCTTCCCGAGGA	2522	
2468	Db	AACGCTGAAATGTACCTAATCTGGTCTCTACTACCAATAGTATGACAGCTTCCCGAGGA	2527	
2523	Qy	TGAATGCTGTTTAAATTTTCAATAAAGTAAATTTGTCACTCTAGCAATTTTGAATGAATAG	2582	
2528	Db	TGAATGCTGTTTAAATTTTCAATAAAGTAAATTTGTCACTCTAGCAATTTTGAATGAATAG	2587	
2583	Qy	TCCTTCACTTTTAAATTTATTCATCTCTCTATTAATAGCATCCCGAGTTTCATGGAGCA	2642	
2588	Db	TCCTTCACTTTTAAATTTATTCATCTCTCTATTAATAGCATCCCGAGTTTCATGGAGCA	2647	
2643	Qy	AAAAACAAGTTTCTGTATTCCTGAAACTTTTCTATGCTCAGTGGAAAGTATCTGCCAGCC	2702	
2648	Db	AAAAACAAGTTTCTGTATTCCTGAAACTTTTCTATGCTCAGTGGAAAGTATCTGCCAGCC	2707	
2703	Qy	ACAGCATGAGGCCTGTGAAGGCTGACTGAGAAATCCCTCTGCTGAAGACCCCTGGTCTCTGT	2762	
2708	Db	ACAGCATGAGGCCTGTGAAGGCTGACTGAGAAATCCCTCTGCTGAAGACCCCTGGTCTCTGT	2767	
2763	Qy	TCCTGCCCTCCAAACATGATATAATTTTATTTGAAATACATATAATCTTTTCTACTATG	2814	
2768	Db	TCCTGCCCTCCAAACATGATATAATTTTATTTGAAATACATATAATCTTTTCTACTATG	2819	

RESULT 5

```

US-10-357-930-25699
; Sequence 25699, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319

```



```
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25699
; LENGTH: 4422
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 4406, 4407, 4408, 4409, 4410, 4411, 4412, 4413, 4414, 4415,
; LOCATION: 4416, 4417, 4418, 4419, 4420, 4421, 4422
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-25699

Query Match          99.0%; Score 2804; DB 9; Length 4422;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2807; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 3 CACGAGCGGGAGTTGGAGCGGATAACGATTTGTGTGTGAGAGGGGCAACGTGCGGATTTTC 62
DB |||
DB 8 CAGTGGCGGGAGTTGGAGCGGATAACGATTTGTGTGTGAGAGGGGCAACGTGCGGATTTTC 67
QY 63 TGTGAACTTTGAGGAGCAATTTCTACGACTTTTCTCTCAGCTGAGGGCTTTTCTCCGACCCCT 122
DB |||
DB 68 TGTGAACTTTGAGGAGCAATTTCTACGACTTTTCTCTCAGCTGAGGGCTTTTCTCCGACCCCT 127
QY 123 GATGCTCTTCAATTCGGTGTCTCGGCGAGCCCGAGCTTGGCGTCTGAGAAATGGATGGTC 182
DB |||
DB 128 GATGCTCTTCAATTCGGTGTCTCGGCGAGCCCGAGCTTGGCGTCTGAGAAATGGATGGTC 187
QY 183 TTCACAAATACCTCTTCAATCCCTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACA 242
DB |||
DB 188 TTCACAAATACCTCTTCAATCCCTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACA 247
QY 243 CACTTCTTATGAGAAACAGGAGTCCAGTTCCTCTCTTTTGGATGPACTTCTCTCTGTC 302
DB |||
DB 248 CACTTCTTATGAGAAACAGGAGTCCAGTTCCTCTCTTTTGGATGPACTTCTCTCTGTC 307
QY 303 TCCCAATATGGAACATGTTACTAGCAGTTGCCAATGAGAAGGCTTTGTTTCGATTTGATATA 362
DB |||
DB 308 TCCCAATATGGAACATGTTACTAGCAGTTGCCAATGAGAAGGCTTTGTTTCGATTTGATATA 367
QY 363 CACAGAAATCACAAGTTTTCAGAAAGAAAGTCTTCAAAGAAATGGATGGCTCACTGGAAATGC 422
DB |||
DB 368 CACAGAAATCACAAGTTTTCAGAAAGAAAGTCTTCAAAGAAATGGATGGCTCACTGGAAATGC 427
QY 423 CGTCTTTGACCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 482
DB |||
DB 428 CGTCTTTGACCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 487
QY 483 AACAGCCAAATTTTGGGACGTAAGCTGGTGGCTGGTGGCTGGTGGCTGGTGGCTGGTGGCTGGCT 542
DB |||
DB 488 AACAGCCAAATTTTGGGACGTAAGCTGGTGGCTGGTGGCTGGTGGCTGGTGGCTGGTGGCTGGCT 547
QY 543 ATGCAGCTTCAAGTCAAGTTCAGTTTCTTAAGTTTGGAAAGCTGTATTTCTGTACGGGTGG 602
DB |||
DB 548 ATGCAGCTTCAAGTCAAGTTCAGTTTCTTAAGTTTGGAAAGCTGTATTTCTGTACGGGTGG 607
QY 603 AAGAGATGGCAACATTTATGCTCTGGATACACAGGTGCAACAAAAGATGGGTTTTATAG 662
DB |||
DB 608 AAGAGATGGCAACATTTATGCTCTGGATACACAGGTGCAACAAAAGATGGGTTTTATAG 667
QY 663 GCNAGTGAATCAATCAGTGGAGCTCACAAATACCTCAGACCAAGCAACCCCTTCAAAACC 722
```

```
DB 668 GCAAGTGAATCAATCAGTGGAGCTCACAATACCTCAGACCAAGCAACCCCTTCAAAACC 727
QY |||
DB 723 CAAGAAGAAACAGAAATTCAGAAAGACTTGTCTCTTGTGTGATTTTCAGCAAGTGTTC 782
DB |||
DB 728 CAAGAAGAAACAGAAATTCAGAAAGACTTGTCTCTTGTGTGATTTTCAGCAAGTGTTC 787
QY |||
DB 783 TGTGGTCTCTTTCAAGACGAGAAATCCTTAGTCTCAGCAGGAGCTGTGGATGGATAT 842
DB |||
DB 788 TGTGGTCTCTTTCAAGACGAGAAATCCTTAGTCTCAGCAGGAGCTGTGGATGGATAT 847
QY |||
DB 843 CAAGATATGGGATTTACGTAAGAAATATATCTGCTTTATCGACAAAGAACCCATAGCATCAA 902
DB |||
DB 848 CAAGATATGGGATTTACGTAAGAAATATATCTGCTTTATCGACAAAGAACCCATAGCATCAA 907
QY |||
DB 903 GTCCTTCTGTACCCAGGTAGCAGCACTCGAAATCTTGGAATATCAAGCTGTGATTTGGA 962
DB |||
DB 908 GTCCTTCTGTACCCAGGTAGCAGCACTCGAAATCTTGGAATATCAAGCTGTGATTTGGA 967
QY |||
DB 963 TTCCACTGGCTCTACTTTTATTTGCTAATTCACAGAGATTAACATCTACATGTTTAAATAT 1022
DB |||
DB 968 TTCCACTGGCTCTACTTTTATTTGCTAATTCACAGAGATTAACATCTACATGTTTAAATAT 1027
QY |||
DB 1023 GACTGGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACCCAGAACTCTACCTTTTA 1082
DB |||
DB 1028 GACTGGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACCCAGAACTCTACCTTTTA 1087
QY |||
DB 1083 TGTAAATCCAGCCTTAGTCCAGATGACAGTGTGTTTGTAGTGTGCTCAAGTGTGAAGC 1142
DB |||
DB 1088 TGTAAATCCAGCCTTAGTCCAGATGACAGTGTGTTTGTAGTGTGCTCAAGTGTGAAGC 1147
QY |||
DB 1143 TGCCTACATATGGAAGGTCTCCACACCTTGGCAACCTCTACTGTGCTCTGGGTCAATTC 1202
DB |||
DB 1148 TGCCTACATATGGAAGGTCTCCACACCTTGGCAACCTCTACTGTGCTCTGGGTCAATTC 1207
QY |||
DB 1203 TCAAGAGTCACTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1262
DB |||
DB 1208 TCAAGAGTCACTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1267
QY |||
DB 1263 TGATGACAATACTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1322
DB |||
DB 1268 TGATGACAATACTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1327
QY |||
DB 1323 TGATGACAATACTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1382
DB |||
DB 1328 TGATGACAATACTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1387
QY |||
DB 1383 AGTAACAGTAAACGAGTAGCAGAGATCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1442
DB |||
DB 1388 AGTAACAGTAAACGAGTAGCAGAGATCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1447
QY |||
DB 1443 ATCCAAATTTTCCCGTCACTCCGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1502
DB |||
DB 1448 ATCCAAATTTTCCCGTCACTCCGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1507
QY |||
DB 1503 TCCTTCAATCTCTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1562
DB |||
DB 1508 TCCTTCAATCTCTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1567
QY |||
DB 1563 CAACAGAGAGGCTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1622
DB |||
DB 1568 CAACAGAGAGGCTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1627
QY |||
DB 1623 GATTAGAAATCTGGGTGAGCCCGAAACACTTCTCTCATCACCCCATCTCTCCACCTGCTTC 1682
DB |||
DB 1628 GATTAGAAATCTGGGTGAGCCCGAAACACTTCTCTCATCACCCCATCTCTCCACCTGCTTC 1687
QY |||
DB 1683 GGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATCTGTGAGCCAGAGATCATCCCA 1742
DB |||
DB 1688 GGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATCTGTGAGCCAGAGATCATCCCA 1747
QY |||
DB 1743 AGCAGAGGCTGTGCTGTGAGTCTTAGAAATAGAGTAAAGAGGAGGCTTAGACTCAAGCTGTCT 1802
```


Db 721 AAAGNACTTGCTCTCTGTGGATTCCAGAAAGTGTACTGTGGTCTCTTTCAAGAC 780
Qy 802 GAGAAATACCTTAGTCTCAGCAGAGCTGTGGATGGGATAATCAAAAGTATGGGATTTACGT 861
Db 781 GAGAAATACCTTAGTCTCAGCAGAGCTGTGGATGGGATAATCAAAAGTATGGGATTTACGT 840
Qy 862 AAGAAATTATACCTTAGTCTCAGCAGAGCCATAGCATCCAAAGTCTTCTGTAGCCAGGT 921
Db 841 AAGAAATTATACCTTAGTCTCAGCAGAAACCCATAGCATCCAAAGTCTTCTGTAGCCAGGT 900
Qy 922 AGCAGCACTCGAAACTCTGGATATTCAAGTCTGATTTTGGATTCACACTGGCTCTACTTTTA 981
Db 901 AGCAGCACTCGAAACTCTGGATATTCAAGTCTGATTTTGGATTCACACTGGCTCTACTTTTA 960
Qy 982 TTTGCTAAATTGCACAGACGATAACTATCATCTATGTTTAAATATGACTGGGTGAAGACTTCT 1041
Db 961 TTTGCTAAATTGCACAGACGATAACTATCATCTATGTTTAAATATGACTGGGTGAAGACTTCT 1020
Qy 1042 CCAGTGGCTATTTCATGGACACCAAGACTCTACCTTTTATGTAATCCAGCCCTTAGT 1101
Db 1021 CCAGTGGCTATTTCATGGACACCAAGACTCTACCTTTTATGTAATCCAGCCCTTAGT 1080
Qy 1102 CCAGATGACAGCTTTTGTAGTCAAGTGGCTCAAGTGTAGTGAAGCTGCCTACATATGGAGGTC 1161
Db 1081 CCAGATGACAGCTTTTGTAGTCAAGTGGCTCAAGTGTAGTGAAGCTGCCTACATATGGAGGTC 1140
Qy 1162 TCCACACCTTGGCAACCTCTACTGTGCTCTGGGTCACTTCTCAAGAGGTCAAGTCTGTG 1221
Db 1141 TCCACACCTTGGCAACCTCTACTGTGCTCTGGGTCACTTCTCAAGAGGTCAAGTCTGTG 1200
Qy 1222 TGGTGTGCTCATCTGACTTTCACAAAGATTGCTACCTGTTCTGTATGACAAATACACTAAAA 1281
Db 1201 TGGTGTGCTCATCTGACTTTCACAAAGATTGCTACCTGTTCTGTATGACAAATACACTAAAA 1260
Qy 1282 ATCTGGCGCTTGAATAGAGCTTTAGAGGAGAAACAGAGAGGTGATAAACTTTCCACGGTG 1341
Db 1261 ATCTGGCGCTTGAATAGAGCTTTAGAGGAGAAACAGAGAGGTGATAAACTTTCCACGGTG 1320
Qy 1342 GGTGTGGCTCTCAGAAAGAAAAAGAGTCAAGACCTTGGCTAGTAAACGATAGCAGAGTAGC 1401
Db 1321 GGTGTGGCTCTCAGAAAGAAAAAGAGTCAAGACCTTGGCTAGTAAACGATAGCAGAGTAGC 1380
Qy 1402 CAGAGTACTCTCCCAAGCCCCCAGGTTAAGTGCATCCATCCAAATCTTCCCGCTCA 1461
Db 1381 CAGAGTACTCTCCCAAGCCCCCAGGTTAAGTGCATCCATCCAAATCTTCCCGCTCA 1440
Qy 1462 TCCGAGCTTGTGCCCAAGCTGTGCTGGAGACTCCCTCTTCTTCCAAATACTCTACG 1521
Db 1441 TCCGAGCTTGTGCCCAAGCTGTGCTGGAGACTCCCTCTTCTTCCAAATACTCTACG 1500
Qy 1522 TTTCTCTATTAAAACTCTCTCGCAAGGCCCGGTCTCCCATCAACAGAAAGAGGCTCTGTC 1581
Db 1501 TTTCTCTATTAAAACTCTCTCGCAAGGCCCGGTCTCCCATCAACAGAAAGAGGCTCTGTC 1560
Qy 1582 TCTCCTGCTCTCCCAAGGCCACTTCATCTTTCAAGATGCGATTAGAACTGGGTGACC 1641
Db 1561 TCTCCTGCTCTCCCAAGGCCACTTCATCTTTCAAGATGCGATTAGAACTGGGTGACC 1620
Qy 1642 CGAACACCTTCTCATCACCCCATCACTCCACCTGTTCCGAGACCAAGATCATGTCT 1701
Db 1621 CGAACACCTTCTCATCACCCCATCACTCCACCTGTTCCGAGACCAAGATCATGTCT 1680
Qy 1702 CCGAGAAAAAGCCCTTATCTCTGTGAGCCAGAAAGTCAATCCCAAGCAGAGGCTTGTCTGTAG 1761
Db 1681 CCGAGAAAAAGCCCTTATCTCTGTGAGCCAGAAAGTCAATCCCAAGCAGAGGCTTGTCTGTAG 1740
Qy 1762 TCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAG 1821
Db 1741 TCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAG 1800
Qy 1822 TGTGTGAAGAGTTGTAACCTGTGTGACTGTAGCTTGTGGCCAAAGTTCGAAATCTTCAATTG 1881
Db 1801 TGTGTGAAGAGTTGTAACCTGTGTGACTGTAGCTTGTGGCCAAAGTTCGAAATCTTCAATTG 1860

Qy 1882 GATCTGTGCTCCCTTGTGTGTAAACAGGAGACCTTAGTAGGACTCTCTAGTCTCCTACC 1941
Db 1861 GATCTGTGCTCCCTTGTGTGTAAACAGGAGACCTTAGTAGGACTCTCTAGTCTCCTACC 1920
Qy 1942 AAATCAAGCAAAATTTGAAGGAGCTGGTACCAAGTATCTCAGAGCCTCGGTCTCTATCAGT 2001
Db 1921 AAATCAAGCAAAATTTGAAGGAGCTGGTACCAAGTATCTCAGAGCCTCGGTCTCTATCAGT 1980
Qy 2002 CCGTATGCTTCAGAAAGCTGTGGAAACGCTACCTCTTCTTTGAGACCTTGTGGAGAAAGG 2061
Db 1981 CCGTATGCTTCAGAAAGCTGTGGAAACGCTACCTCTTCTTTGAGACCTTGTGGAGAAAGG 2040
Qy 2062 TCTGAATGTTAGCAAGAGAGATAGTTCGCCAGAGATAAACTGGTCTTGGCCATG 2121
Db 2041 TCTGAATGTTAGCAAGAGAGATAGTTCGCCAGAGATAAACTGGTCTTGGCCATG 2100
Qy 2122 GCAGCAAAACGGAAGGCTGAGAATCCATCTCCAGAAAGTCCGTCTATCCAGACACCCCAAT 2181
Db 2101 GCAGCAAAACGGAAGGCTGAGAATCCATCTCCAGAAAGTCCGTCTATCCAGACACCCCAAT 2160
Qy 2182 TCCAGGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTCAACATCAAGCCAGGTCTCCATG 2241
Db 2161 TCCAGGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTCAACATCAAGCCAGGTCTCCATG 2220
Qy 2242 AGGAAATCTGCACATATCTCCATAGAAAGTCCAGGAGGACTTCTGTGGTCTTGACAC 2301
Db 2221 AGGAAATCTGCACATATCTCCATAGAAAGTCCAGGAGGACTTCTGTGGTCTTGACAC 2280
Qy 2302 TCAACAGAAATTTATAGATTCTTAATCTGAGTGTAGTTACTGAGCTTTGGTCCACTAAAAACAAG 2361
Db 2281 TCAACAGAAATTTATAGATTCTTAATCTGAGTGTAGTTACTGAGCTTTGGTCCACTAAAAACAAG 2340
Qy 2362 CTGAGCTTTGGTCCACTAAAAACAAGATGAAAAATAACAAGAGTGAATCTATAACTCTGTC 2421
Db 2341 CTGAGCTTTGGTCCACTAAAAACAAGATGAAAAATAACAAGAGTGAATCTATAACTCTGTC 2400
Qy 2422 TTTAAGAAAGCTGCTTTTCAATTTTAGACAAAAATCTTTTCAACGCTGAAATGTACCTAA 2481
Db 2401 TTTAAGAAAGCTGCTTTTCAATTTTAGACAAAAATCTTTTCAACGCTGAAATGTACCTAA 2460
Qy 2482 TCTGTGTTCTACTACCATATGTATATGAGCTTCCAGAGGATGAATGCTGTGTTTAAAT 2541
Db 2461 TCTGTGTTCTACTACCATATGTATATGAGCTTCCAGAGGATGAATGCTGTGTTTAAAT 2520
Qy 2542 TCATAAAGTAAATTTGTCACTCTAGCATTTTGAATGAATGTCTTCACTTTTAAATTTAT 2601
Db 2521 TCATAAAGTAAATTTGTCACTCTAGCATTTTGAATGAATGTCTTCACTTTTAAATTTAT 2580
Qy 2602 TCATCTTCTCTATAATATGACATCCAGTTTCATGGAGGCAAAAACAAGTTTCTTGTTA 2661
Db 2581 TCATCTTCTCTATAATATGACATCCAGTTTCATGGAGGCAAAAACAAGTTTCTTGTTA 2640
Qy 2662 TCCTGAAACTTTCTATGCTCAGTGTGAAAGTATCTGCCAGCCACACGATGAGGCTGTGAA 2721
Db 2641 TCCTGAAACTTTCTATGCTCAGTGTGAAAGTATCTGCCAGCCACACGATGAGGCTGTGAA 2700
Qy 2722 GGCTGACTGAGAAATCTCTCTGCTGAAGACCCCTGGTCTGTGCTCCCAACATGTATA 2781
Db 2701 GGCTGACTGAGAAATCTCTCTGCTGAAGACCCCTGGTCTGTGCTCCCAACATGTATA 2760
Qy 2782 ATTTTATTTGAATACATAATCTTTTCACTATG 2814
Db 2761 ATTTTATTTGAATACATAATCTTTTCACTATG 2793

RESULT 7

US-10-393-590-85
; Sequence 85, Application US/10393590
; Publication No. US20030190656A1
; GENERAL INFORMATION:
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: BREAST CANCER PROGNASTIC PORTFOLIO

FILE REFERENCE: CDS 268 US NP
CURRENT APPLICATION NUMBER: US/10/393,590
CURRENT FILING DATE: 2003-03-21
PRIOR APPLICATION NUMBER: 60/368,789
PRIOR FILING DATE: 2002-03-29
NUMBER OF SEQ ID NOS: 100
SOFTWARE: PatentIn version 3.1
SEQ ID NO 85
LENGTH: 4221
TYPE: DNA
ORGANISM: human
us-10-393-590-85

Query Match 98.5%; Score 2789,8; DB 7; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	22	CGATAACGATTTGGTTGTGAGAGGCGCAACGTCGCAGTTCTCTGCAACTTGGAGCAAT	81
DB	1	CGATAACGATTTGGTTGTGAGAGGCGCAACGTCGCAGTTCTCTGCAACTTGGAGCAAT	60
QY	82	TCTACGACTTTCTCTCAGCTGAGGCTTTTCCTCCGACCCCTGATGCTCTCAATTCGGTG	141
DB	61	TCTACGACTTTCTCTCAGCTGAGGCTTTTCCTCCGACCCCTGATGCTCTCAATTCGGTG	120
QY	142	CTCGCCAGCCCGCAGCTTGGCGCTCTGAGAAATGGATGGTCTTTCACAAATACCCCTCTTCAA	201
DB	121	CTCGCCAGCCCGCAGCTTGGCGCTCTGAGAAATGGATGGTCTTTCACAAATACCCCTCTTCAA	180
QY	202	TCCCTTTCTGACTGTTATCAGTCAGTCAGTGGTAAATGATGAACACACTTTCTTATGGAGAAACA	261
DB	181	TCCCTTTCTGACTGTTATCAGTCAGTCAGTGGTAAATGATGAACACACTTTCTTATGGAGAAACA	240
QY	262	GGAGTCCAGTTCCTCTTTTGGATGATACCTTCTCTCTGCTCCCAATATGGAACATGTA	321
DB	241	GGAGTCCAGTTCCTCTTTTGGATGATACCTTCTCTCTGCTCCCAATATGGAACATGTA	300
QY	322	CTAGCAGTTTCCCAATGAAGAAGCTTTGTCGATTTGATATAACACAGATCACAAGATTTTC	381
DB	301	CTAGCAGTTTCCCAATGAAGAAGCTTTGTCGATTTGATATAACACAGATCACAAGATTTTC	360
QY	382	AGAAAGAGTTCCTTCAAGAAATGGATGGCTCACTGGAAATGCCGTCTTTTGACCTGGCGTGG	441
DB	361	AGAAAGAGTTCCTTCAAGAAATGGATGGCTCACTGGAAATGCCGTCTTTTGACCTGGCGTGG	420
QY	442	GTTCTGGTGAACTTAACTTTGTTACAGCAGCAGGTGATCAAAACAGCAAAATTTTGGGAC	501
DB	421	GTTCTGGTGAACTTAACTTTGTTACAGCAGCAGGTGATCAAAACAGCAAAATTTTGGGAC	480
QY	502	GTAAGAGTGGTGAGCTGATTGGAAACATGCAGAGGTCAATCAATGCAGCCTCAAGTCAGTT	561
DB	481	GTAAGAGTGGTGAGCTGATTGGAAACATGCAGAGGTCAATCAATGCAGCCTCAAGTCAGTT	540
QY	562	GCCTTTTCTAAGTTTGAAGAGCTGATTCTGTACGGGTGGAAAGAGATGGCAACATTATG	621
DB	541	GCCTTTTCTAAGTTTGAAGAGCTGATTCTGTACGGGTGGAAAGAGATGGCAACATTATG	600
QY	622	GTCCTGGGATACCGAGTGCAACAAAAGATGGGTTTATAGGCAAGTGGAATCAAAATCAGT	681
DB	601	GTCCTGGGATACCGAGTGCAACAAAAGATGGGTTTATAGGCAAGTGGAATCAAAATCAGT	660
QY	682	GGAGCTCACAAATACCTCAGACAGCAACCCCTTCAAAACCCCAAGAGAAACAGAAATTC	741
DB	661	GGAGCTCACAAATACCTCAGACAGCAACCCCTTCAAAACCCCAAGAGAAACAGAAATTC	720
QY	742	AAAGGACTTCTCTCTGTGGATTTCCAGCAAGTGTACTGTGGTCTCTTTTCAAGAC	801
DB	721	AAAGGACTTCTCTCTGTGGATTTCCAGCAAGTGTACTGTGGTCTCTTTTCAAGAC	780
QY	802	GAGAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAGATATGGGATTTACGT	861
DB	781	GAGAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAGATATGGGATTTACGT	840

QY	862	AAGAATTATATGCTTTATCGACAAGAACCCATAGATCAAGTCTTTCTGTATCCAGGT	921
DB	841	AAGAATTATATGCTTTATCGACAAGAACCCATAGATCAAGTCTTTCTGTATCCAGGT	900
QY	922	AGCAGACCTCGAAAACCTTGGATATTCAAGTCTGATTTTGGATTCACCTGGCTCTACTTTA	981
DB	901	AGCAGACCTCGAAAACCTTGGATATTCAAGTCTGATTTTGGATTCACCTGGCTCTACTTTA	960
QY	982	TTTGTCTAATTGCACAGACGATACATCTACATGTTTAAATATGACTGGGTGAAGACTTCT	1041
DB	961	TTTGTCTAATTGCACAGACGATACATCTACATGTTTAAATATGACTGGGTGAAGACTTCT	1020
QY	1042	CCAGTGGCTATTTTCAATGGACACACAGAACTCTACCTTTTATGTAAATCCAGCCTTAGT	1101
DB	1021	CCAGTGGCTATTTTCAATGGACACACAGAACTCTACCTTTTATGTAAATCCAGCCTTAGT	1080
QY	1102	CCAGATGACACAGTTTTTTAGTCACTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT	1161
DB	1081	CCAGATGACACAGTTTTTTAGTCACTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT	1140
QY	1162	TCCACACCTCGCAACCTCTACTGCTCTCTGGGTCACTTCTCAAGAGGTCAAGTCAAGT	1221
DB	1141	TCCACACCTCGCAACCTCTACTGCTCTCTGGGTCACTTCTCAAGAGGTCAAGTCAAGT	1200
QY	1222	TGCTGCTGCTCACTGACTTCAAAAGATTTGCTACCTGTTCTGATGACAATACACTAAAA	1281
DB	1201	TGCTGCTGCTCACTGACTTCAAAAGATTTGCTACCTGTTCTGATGACAATACACTAAAA	1260
QY	1282	ATCTGCGCTTGAATAGAGGCTTAGAGGAGAAACAGAGAGGTGATATAACTTTCCACGGTG	1341
DB	1261	ATCTGCGCTTGAATAGAGGCTTAGAGGAGAAACAGAGAGGTGATATAACTTTCCACGGTG	1320
QY	1342	GGTGGGCTCTCAGAGAGAAAGAGTCAAGACCTGGGCTAGTAAACAGTAAAGAGTAGC	1401
DB	1321	GGTGGGCTCTCAGAGAGAAAGAGTCAAGACCTGGGCTAGTAAACAGTAAAGAGTAGC	1380
QY	1402	CAGAGTACTCTGCTCCAAAGCCCCCAGGGTAAAGTGAATCCATCCAAATTTTCCCCGTCA	1461
DB	1381	CAGAGTACTCTGCTCCAAAGCCCCCAGGGTAAAGTGAATCCATCCAAATTTTCCCCGTCA	1440
QY	1462	TCCGAGCTTGTGCCCCAAGCTGTGTGGAGACCTCCCTCTTCTTCAAAATCTCTCTACG	1521
DB	1441	TCCGAGCTTGTGCCCCAAGCTGTGTGGAGACCTCCCTCTTCTTCAAAATCTCTCTACG	1500
QY	1522	TTTCTCTTAAACCTCTCTGCGCAGGCCCGGTCTCCCATCAACAGAGAGGCTCTGTTC	1581
DB	1501	TTTCTCTTAAACCTCTCTGCGCAGGCCCGGTCTCCCATCAACAGAGAGGCTCTGTTC	1560
QY	1582	TCTCCGTCTCTCCCAAGCCACCTTTCATCTTTCAAGATGTCGATTAGAAACTGGGTGACC	1641
DB	1561	TCTCCGTCTCTCCCAAGCCACCTTTCATCTTTCAAGATGTCGATTAGAAACTGGGTGACC	1620
QY	1642	CGAACACCTTCTCTATCACACCCATCTCACTCACTGCTTGGAGACCAAGATCATGTCT	1701
DB	1621	CGAACACCTTCTCTATCACACCCATCTCACTCACTGCTTGGAGACCAAGATCATGTCT	1680
QY	1702	CCGAGAAAGCCCTTATTTCTGTCGACCAAGAGTCAATCCCAAGCAGAGGCTTGTCTGAG	1761
DB	1681	CCGAGAAAGCCCTTATTTCTGTCGACCAAGAGTCAATCCCAAGCAGAGGCTTGTCTGAG	1740
QY	1762	TCTAGAAATAGAGTAAAGAGGAGGCTTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAAG	1821
DB	1741	TCTAGAAATAGAGTAAAGAGGAGGCTTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAAG	1800
QY	1822	TGTGTGAAGAGTTGAACTGTGTGAGTGAAGTGGCCAAAGTTGAAATCTTCAATTTG	1881
DB	1801	TGTGTGAAGAGTTGAACTGTGTGAGTGAAGTGGCCAAAGTTGAAATCTTCAATTTG	1860
QY	1882	GATCTGTGCTGCTGCTGTTAACCCAGGAGACCTTAGTAGGACTCTCTAGTCTCTACC	1941
DB	1861	GATCTGTGCTGCTGCTGTTAACCCAGGAGACCTTAGTAGGACTCTCTAGTCTCTACC	1920
QY	1942	AAATCAAGCAAAATTTGAAGAGCTGGTACCAGTATCTCTCAGAGCCTCCGCTCTCTATCAGT	2001

```
Db 1921 AAATCAAGCAAAATTGAGAGAGCTGTACAGTAATCTCAGAGCCTCCGCTCTCTATCAGT 1980
Qy 2002 CCGTATGCTTCAGAAAGCTGTGGAACGCTACCTCTTCCTTTGAGAGCTTTGTGGAGAAGG 2061
Db 1981 CCGTATGCTTCAGAAAGCTGTGGAACGCTACCTCTTCCTTTGAGAGCTTTGTGGAGAAGG 2040
Qy 2062 TCTGAAATGGTAGGCAAGAGAAATAGTTCCTCCAGAGAAATAAATAAACHGGTTGTGGCCATG 2121
Db 2041 TCTGAAATGGTAGGCAAGAGAAATAGTTCCTCCAGAGAAATAAATAAACHGGTTGTGGCCATG 2100
Qy 2122 GCAGCCAAACGGGAAGCTGAGAAATCATCTCCACGAAGTCCGTCATCCAGACACCCCAAT 2181
Db 2101 GCAGCCAAACGGGAAGCTGAGAAATCATCTCCACGAAGTCCGTCATCCAGACACCCCAAT 2160
Qy 2182 TCCAGGAGACAGACGCGGAAAGACATTTGCCAAGCCCGGTCAACATCAGCCAGCTCCATG 2241
Db 2161 TCCAGGAGACAGACGCGGAAAGACATTTGCCAAGCCCGGTCAACATCAGCCAGCTCCATG 2220
Qy 2242 AGGAAATCTGCACATACCTTCCATAGAAAGTCCAGGAGACTTCTGTGGTCTCTGAACAC 2301
Db 2221 AGGAAATCTGCACATACCTTCCATAGAAAGTCCAGGAGACTTCTGTGGTCTCTGAACAC 2280
Qy 2302 TCAACAGAAATTATAGATTCTAATCTCAGTGAGTTACTGAGCTTTGGTCCACTAATAACACAG 2361
Db 2281 TCAACAGAAATTATAGATTCTAATCTCAGTGAGTTACTGAGCTTTGGTCCACTAATAACACAG 2340
Qy 2362 CTGAGCTTTGGTCCACTAATAACAGATGAAATAACAAGAGTGACTCTATAACTCTGGTC 2421
Db 2341 CTGAGCTTTGGTCCACTAATAACAGATGAAATAACAAGAGTGACTCTATAACTCTGGTC 2400
Qy 2422 TTTAAGAAAGCTGCCCTTTTCATTTTATAGACAAATCTTTTCAACGCTGAAATGTACTTAA 2481
Db 2401 TTTAAGAAAGCTGCCCTTTTCATTTTATAGACAAATCTTTTCAACGCTGAAATGTACTTAA 2460
Qy 2482 TCTGGTTCTACTACCAATAATGTATATATGACAGTTCCTCCGAGGAGTGAATGCTGTAAAT 2541
Db 2461 TCTGGTTCTACTACCAATAATGTATATATGACAGTTCCTCCGAGGAGTGAATGCTGTAAAT 2520
Qy 2542 TCATAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTAAT 2601
Db 2521 TCATAAGTAAATTTGTCACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTAAT 2580
Qy 2602 TCATCTCTCTATAATAATGACATCCAGTTTCATGGAGGCAAAAAACAAGTTTCTTGTTA 2661
Db 2581 TCATCTCTCTATAATAATGACATCCAGTTTCATGGAGGCAAAAAACAAGTTTCTTGTTA 2640
Qy 2662 TCCGTAACCTTTCTATGCTCAGTGGAAAGTATCTGCCAGCCAGCATGAGGCTGTGAA 2721
Db 2641 TCCGTAACCTTTCTATGCTCAGTGGAAAGTATCTGCCAGCCAGCATGAGGCTGTGAA 2700
Qy 2722 GGCTGACTGAGAAATCCTCTGCTGGAAGACCCCTGGTTCTGTTCTGCTCCAAATGATA 2781
Db 2701 GGCTGACTGAGAAATCCTCTGCTGGAAGACCCCTGGTTCTGTTCTGCTCCAAATGATA 2760
Qy 2782 ATTTTATTTGAAATACATAATCTTTTCACTATG 2814
Db 2761 ATTTTATTTGAAATACATAATCTTTTCACTATG 2793
```

RESULT 8

US-10-393-567-85

; Sequence 85, Application US/10393567

; Publication No. US20030194733A1

; GENERAL INFORMATION:

; APPLICANT: WANG, YIXIN

; TITLE OF INVENTION: CANCER DIAGNOSTIC PANEL

; FILE REFERENCE: CDS 269 US NP

; CURRENT APPLICATION NUMBER: US/10393,567

; PRIOR FILING DATE: 2003-03-21

; PRIOR APPLICATION NUMBER: 60/368,667

; PRIOR FILING DATE: 2002-03-29

; NUMBER OF SEQ ID NOS: 100

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 85

; LENGTH: 4221

; TYPE: DNA

; ORGANISM: human

US-10-393-567-85

Query Match 98.5%; Score 2789.8; DB 7; Length 4221;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy 22 CGATAACCATTTGTCTGTGAGAGCGCAACGTGCGATTTCTGCTGTAACCTTTGGAGGCAAT 81
Db 1 COATHAACCATTTGTCTGTGAGAGCGCAACGTGCGATTTCTGCTGTAACCTTTGGAGGCAAT 60
Qy 82 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG 141
Db 61 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG 120
Qy 142 CTCGCCACGCCAGCTTTGGGCTCTGAGAAATGGATGGTCTTACAAATACCCCTTTCAA 201
Db 121 CTCGCCACGCCAGCTTTGGGCTCTGAGAAATGGATGGTCTTACAAATACCCCTTTCAA 180
Qy 202 TCCCTTCTGACTGGTTATCAGTGCAAGTGAATGAACACACTTCTTATGGAGAAACA 261
Db 181 TCCCTTCTGACTGGTTATCAGTGCAAGTGAATGAACACACTTCTTATGGAGAAACA 240
Qy 262 GGAGTCCCAGTTCTCCCTTTTGGATGTACCTTCTCTGTCTCCCAATATGGAACATGTA 321
Db 241 GGAGTCCCAGTTCTCCCTTTTGGATGTACCTTCTCTGTCTCCCAATATGGAACATGTA 300
Qy 322 CTAGCAGTTGCCAATGAAGAAGGCTTTGTTGATTTGATAACACAGAAATCAAAAGTTTC 381
Db 301 CTAGCAGTTGCCAATGAAGAAGGCTTTGTTGATTTGATAACACAGAAATCAAAAGTTTC 360
Qy 382 AGAAAGAAAGTGCTTCAAGAAATGGAATGGTCACTGGAATCGCGTCTTGGACCTGCGCTGG 441
Db 361 AGAAAGAAAGTGCTTCAAGAAATGGAATGGTCACTGGAATCGCGTCTTGGACCTGCGCTGG 420
Qy 442 GTTCTCGTGAACCTTAAACTTTGTTACAGCAGCAGGTGATCAAAACAGCCAAATTTGGGAC 501
Db 421 GTTCTCGTGAACCTTAAACTTTGTTACAGCAGCAGGTGATCAAAACAGCCAAATTTGGGAC 480
Qy 502 GTAAAAGCTGTGAGCTGATTTGGAAACATGCAAAAGTCAATCAATGAGCCTCAAGTCAGTT 561
Db 481 GTAAAAGCTGTGAGCTGATTTGGAAACATGCAAAAGTCAATCAATGAGCCTCAAGTCAGTT 540
Qy 562 GCCTTTTCTAAGTTTGAGAAAGCTGTATTCTGTACGGGTGGAAAGATGGCAACATTATG 621
Db 541 GCCTTTTCTAAGTTTGAGAAAGCTGTATTCTGTACGGGTGGAAAGATGGCAACATTATG 600
Qy 622 GTCTGGGATACAGGTGCAACAAAAGATGGGTTTTATAGGCAAGTGAATCAAAATCAGT 681
Db 601 GTCTGGGATACAGGTGCAACAAAAGATGGGTTTTATAGGCAAGTGAATCAAAATCAGT 660
Qy 682 GGAGCTCACAATACCTCAGACAAAGCAAAACCCCTTCAAAACCCCAAGAGAAACAGAAATCA 741
Db 661 GGAGCTCACAATACCTCAGACAAAGCAAAACCCCTTCAAAACCCCAAGAGAAACAGAAATCA 720
Qy 742 AAAGGACTTGTCTCTGTGGATTTCCAGCAAAAGTGTACTGTGTGCTCTCTTTCAAGAC 801
Db 721 AAAGGACTTGTCTCTGTGGATTTCCAGCAAAAGTGTACTGTGTGCTCTCTTTCAAGAC 780
Qy 802 GAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAAGTATGGGATTTACGT 861
Db 781 GAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAAGTATGGGATTTACGT 840
Qy 862 AAGAATTATCTGCTTATCGACAAGAACCCATAGCATCCAGTCTTCTCTGTACCCAGT 921
Db 841 AAGAATTATCTGCTTATCGACAAGAACCCATAGCATCCAGTCTTCTCTGTACCCAGT 900
Qy 922 AGCAGCACTCGAAAAACTTGGGATATTCAAGTCTGTGATTTTGGATTTCCACTGGCTCTACTTTA 981
```

Db 901 AGCAGCACTCGAAAACTTGATATTCAGAGTCTGATTTTGGATTCCTCACTGGCTCTACTTTA 960
Qy 982 TTTTGTAAATTCACAGCAGATACATCTACATCTTTAAATATGACTGGTGTGAGACTTCT 1041
Db 961 TTTTGTAAATTCACAGCAGATACATCTACATCTTTAAATATGACTGGTGTGAGACTTCT 1020
Qy 1042 CCAGTGGCTATTTTCAATGAGCACCAGAACTCTACCTTTTATGTAAATTCACGCTTAGT 1101
Db 1021 CCAGTGGCTATTTTCAATGAGCACCAGAACTCTACCTTTTATGTAAATTCACGCTTAGT 1080
Qy 1102 CCAGATGACCAAGTTTATGCTAGTGGCTCAAGTATGATGAAGTCCCTACATATGGAAGGTC 1161
Db 1081 CCAGATGACCAAGTTTATGCTAGTGGCTCAAGTATGATGAAGTCCCTACATATGGAAGGTC 1140
Qy 1162 TCCACACCTCGGCAACTCTGCTCTGCTCTCGGTCATTTCTCAAGAGGTCAGCTCTGTG 1221
Db 1141 TCCACACCTCGGCAACTCTGCTCTGCTCTCGGTCATTTCTCAAGAGGTCAGCTCTGTG 1200
Qy 1222 TGCTGGTGTCCATCTGACTTTCACAAAGATTGCTACCTGTTCTGATGACAAATACACTAAAA 1281
Db 1201 TGCTGGTGTCCATCTGACTTTCACAAAGATTGCTACCTGTTCTGATGACAAATACACTAAAA 1260
Qy 1282 ATCTGGCTGTGAATGAGGCTTAGAGGAGAAACCAAGAGGTTGATAACTTTCCACGGTG 1341
Db 1261 ATCTGGCTGTGAATGAGGCTTAGAGGAGAAACCAAGAGGTTGATAACTTTCCACGGTG 1320
Qy 1342 GGTGGGCTCTCAGAGAAAAGAGTCAAGACCTGGCTAGTACAGTAAGCAGTAGC 1401
Db 1321 GGTGGGCTCTCAGAGAAAAGAGTCAAGACCTGGCTAGTACAGTAAGCAGTAGC 1380
Qy 1402 CAGAGTACTCTCCGCAAGCCCGAGGTAAGTGCAATCCATCCAAATTTCTCCCGTCA 1461
Db 1381 CAGAGTACTCTCCGCAAGCCCGAGGTAAGTGCAATCCATCCAAATTTCTCCCGTCA 1440
Qy 1462 TCCGAGCTGTGCCCCCAAGCTGTGTGGAGACTCCCTCTTCCCTTCAAACTCTCTACG 1521
Db 1441 TCCGAGCTGTGCCCCCAAGCTGTGTGGAGACTCCCTCTTCCCTTCAAACTCTCTACG 1500
Qy 1522 TTCTCTATTAAACCTCTCTGSCAAGGCCCGCTCTCCCATCAAGAGAGGCTCTGTG 1581
Db 1501 TTCTCTATTAAACCTCTCTGSCAAGGCCCGCTCTCCCATCAAGAGAGGCTCTGTG 1560
Qy 1582 TCCCTCGTCTCTCCCAAGCCACTTCATCTTCAAGATGCGATTAGAACTGGGTGACC 1641
Db 1561 TCCCTCGTCTCTCCCAAGCCACTTCATCTTCAAGATGCGATTAGAACTGGGTGACC 1620
Qy 1642 CGAACACCTTCTCATCACACCCATCACTCCACCTGCTTCCGAGACCAAGATCATGTCT 1701
Db 1621 CGAACACCTTCTCATCACACCCATCACTCCACCTGCTTCCGAGACCAAGATCATGTCT 1680
Qy 1702 CCGAGAAAAGCCCTTATCTGTGAGCCAGAGTCACTCCAGCAGAGGCTTGTCTGAG 1761
Db 1681 CCGAGAAAAGCCCTTATCTGTGAGCCAGAGTCACTCCAGCAGAGGCTTGTCTGAG 1740
Qy 1762 TCTAGAAAATAGATGAAAGAGGCTTAGACTCAAGCTGTCTGAGAGTGTGAAACAAAAG 1821
Db 1741 TCTAGAAAATAGATGAAAGAGGCTTAGACTCAAGCTGTCTGAGAGTGTGAAACAAAAG 1800
Qy 1822 TGTGTGAAGAGTTGTAACCTGTGTGACTGAGCTTGTGAGGCAAGTGTGAAATCTTCAATTG 1881
Db 1801 TGTGTGAAGAGTTGTAACCTGTGTGACTGAGCTTGTGAGGCAAGTGTGAAATCTTCAATTG 1860
Qy 1882 GATCTGTGCTGCTTGTGTGTAAACAGGAGACTTGTAGTAAGACTCTCTAGGTCTTACC 1941
Db 1861 GATCTGTGCTGCTTGTGTGTAAACAGGAGACTTGTAGTAAGACTCTCTAGGTCTTACC 1920
Qy 1942 AATCAAGCAAAATTCAGAGGCTGTGACAGTATCTCAGAGCTCCGCTCTCTATCAGT 2001
Db 1921 AATCAAGCAAAATTCAGAGGCTGTGACAGTATCTCAGAGCTCCGCTCTCTATCAGT 1980
Qy 2002 CCGTATGCTTTCAGAAAGCTGTGAAACGCTACCTCTTCTTTCAGAGCTTGTGGAGAGGG 2061
Db 1981 CCGTATGCTTTCAGAAAGCTGTGAAACGCTACCTCTTCTTTCAGAGCTTGTGGAGAGGG 2040

Qy 2062 TCTGAAATCGTAGGCAAGAGAAATAGTTCCCGAGAGAAATAAACTGGTTGTGGCCATG 2121
Db 2041 TCTGAAATCGTAGGCAAGAGAAATAGTTCCCGAGAGAAATAAACTGGTTGTGGCCATG 2100
Qy 2122 GCAGCAAAACGGAAGCTGAGAAATCACTCTCAAGAGTCCGTCATCCAGACACCCAAAT 2181
Db 2101 GCAGCAAAACGGAAGCTGAGAAATCACTCTCAAGAGTCCGTCATCCAGACACCCAAAT 2160
Qy 2182 TCCAGGAGACAGCGGGAAGACATTCGCAAGCGGTCACCATCAAGCCAGCTCCATG 2241
Db 2161 TCCAGGAGACAGCGGGAAGACATTCGCAAGCGGTCACCATCAAGCCAGCTCCATG 2220
Qy 2242 AGCAAAATCTGCACATACCTTCCATAGAAAGTCCAGGAGGACTTCTGTGCTCTGAACAC 2301
Db 2221 AGCAAAATCTGCACATACCTTCCATAGAAAGTCCAGGAGGACTTCTGTGCTCTGAACAC 2280
Qy 2302 TCAACAGAAATATAGATCTAATCTGAGTGAATTAAGAGTCTCTATAACTCTGGTC 2361
Db 2281 TCAACAGAAATATAGATCTAATCTGAGTGAATTAAGAGTCTCTATAACTCTGGTC 2340
Qy 2362 CTGAGCTTTGGTCCACTAAACAGATGAAATTAAGAGTCTCTATAACTCTGGTC 2421
Db 2341 CTGAGCTTTGGTCCACTAAACAGATGAAATTAAGAGTCTCTATAACTCTGGTC 2400
Qy 2422 TTTTAAAGAGCTGCTTTTTCATTTTGTAGCAAAATCTTTTCAAGCTGAAATGTACTAA 2481
Db 2401 TTTTAAAGAGCTGCTTTTTCATTTTGTAGCAAAATCTTTTCAAGCTGAAATGTACTAA 2460
Qy 2482 TCTGGTCTTCTACTACCATATATATGATGACAGTCCGAGGATGAATGTGTGTTAAAT 2541
Db 2461 TCTGGTCTTCTACTACCATATATATGATGACAGTCCGAGGATGAATGTGTGTTAAAT 2520
Qy 2542 TCATAAGTAAATTTGTCTACTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTAAT 2601
Db 2521 TCATAAGTAAATTTGTCTACTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTAAT 2580
Qy 2602 TCATCTTCTCTATAATATGATGACATCCAGTTCATGAGGCAAAACAAAGTTTCTGTGA 2661
Db 2581 TCATCTTCTCTATAATATGATGACATCCAGTTCATGAGGCAAAACAAAGTTTCTGTGA 2640
Qy 2662 TCCGAAACTTTCTATGCTCAGTGGAAAGTATCTGCCAGCCACAGCATGAGGCTGTGAA 2721
Db 2641 TCCGAAACTTTCTATGCTCAGTGGAAAGTATCTGCCAGCCACAGCATGAGGCTGTGAA 2700
Qy 2722 GGCTGACTGAGAAATCTCTGCTGAGAGCCCTGTTCTGTTCTGCTTCCACATGTATA 2781
Db 2701 GGCTGACTGAGAAATCTCTGCTGAGAGCCCTGTTCTGTTCTGCTTCCACATGTATA 2760
Qy 2782 ATTTTATTTGAAATACATAATCTTTTCACTATG 2814
Db 2761 ATTTTATTTGAAATACATAATCTTTTCACTATG 2793

RESULT 9

US-10-394-087-85
; Sequence 85, Application US/10394087
; Publication NO. US20030194734A1
; GENERAL INFORMATION:
; APPLICANT: Jatkoe, Tim
; TITLE OF INVENTION: SELECTION OF MARKERS
; FILE REFERENCE: CDS 265 US NP
; CURRENT APPLICATION NUMBER: US/10/394,087
; PRIOR FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: 60/368,790
; PRIOR FILING DATE: 2002-03-29
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 85
; LENGTH: 4221
; TYPE: DNA
; ORGANISM: human
US-10-394-087-85

Query Match				98.5%; Score 2789.8; DB 7; Length 4221;			
Best Local Similarity 99.9%; Pred. No. 0;							
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;							
Qy	22	CGATAACGATTTGTGTGAGAGCGCAACGTGCGATTTCTGCTGAACCTTGGAGGCATT	81				
Db	1	CGATAACGATTTGTGTGAGAGCGCAACGTGCGATTTCTGCTGAACCTTGGAGGCATT	60				
Qy	82	TCTACGACTTTCTCTCAGCTGAGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG	141				
Db	61	TCTACGACTTTCTCTCAGCTGAGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG	120				
Qy	142	CTCCGCCAGCCCGCTGGCGCTCTGAGAAATGATGGTCTTCAATACCCCTCTTCAA	201				
Db	121	CTCCGCCAGCCCGCTGGCGCTCTGAGAAATGATGGTCTTCAATACCCCTCTTCAA	180				
Qy	202	TCCCTTCTGACTGGTTATCAGTGCAGTGGTAATGATGAACACACTTCTTATGGAGAAACA	261				
Db	181	TCCCTTCTGACTGGTTATCAGTGCAGTGGTAATGATGAACACACTTCTTATGGAGAAACA	240				
Qy	262	GGAGTCCCAAGTTCCTCTTTTGGATGTACCTTCTCTTCTGCTCCCAATATGGAACATGTA	321				
Db	241	GGAGTCCCAAGTTCCTCTTTTGGATGTACCTTCTCTTCTGCTCCCAATATGGAACATGTA	300				
Qy	322	CTAGCAGTTCCTCAATGAAGAGCTTTGTTCGATTTGTAATACACAGATTCACAAAGTTTC	381				
Db	301	CTAGCAGTTCCTCAATGAAGAGCTTTGTTCGATTTGTAATACACAGATTCACAAAGTTTC	360				
Qy	382	AGAAAGAGTGTCTCAAGAAATGGATGGCTCACTCGAATGCCGTCTTGGACCTGCGCTGG	441				
Db	361	AGAAAGAGTGTCTCAAGAAATGGATGGCTCACTCGAATGCCGTCTTGGACCTGCGCTGG	420				
Qy	442	GTTCTCTGTGTAACTTAAACTTGTACAGCAGCAGGTGATCAACAGCCAAATTTTGGGAC	501				
Db	421	GTTCTCTGTGTAACTTAAACTTGTACAGCAGCAGGTGATCAACAGCCAAATTTTGGGAC	480				
Qy	502	GTAAGCTGTGTGAGTGAATGGAACATGCCAAGGTCAATGAGCCCTCAAGTCAGTT	561				
Db	481	GTAAGCTGTGTGAGTGAATGGAACATGCCAAGGTCAATGAGCCCTCAAGTCAGTT	540				
Qy	562	GCTTTTCTAAAGTTTGAAGAGCTGTATCTGTACGGTGGAGAGATGGCAACATTATG	621				
Db	541	GCTTTTCTAAAGTTTGAAGAGCTGTATCTGTACGGTGGAGAGATGGCAACATTATG	600				
Qy	622	GTCTGGGATACAGGTGCAACAAAAGATGGTTTTATAGCAAGTGAATCAAAATCAGT	681				
Db	601	GTCTGGGATACAGGTGCAACAAAAGATGGTTTTATAGCAAGTGAATCAAAATCAGT	660				
Qy	682	GGAGTCACAATACCTCAGACAGCAACCCCTTCAAAACCCCAAGCAAGCAAGATTC	741				
Db	661	GGAGTCACAATACCTCAGACAGCAACCCCTTCAAAACCCCAAGCAAGCAAGATTC	720				
Qy	742	AAAGGACTTGTCTCTGTGATTTCCAGCAAGTGTTCAGTGTCTCTTTTCAAGAC	801				
Db	721	AAAGGACTTGTCTCTGTGATTTCCAGCAAGTGTTCAGTGTCTCTTTTCAAGAC	780				
Qy	802	GAGAAATACCTTGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAAGTATGGGATTTACGT	861				
Db	781	GAGAAATACCTTGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAAGTATGGGATTTACGT	840				
Qy	862	AAGAAATATACCTTGTCTCAGCAAGAACCCATAGCATCCAGTCTTCTGTACCCAGGT	921				
Db	841	AAGAAATATACCTTGTCTCAGCAAGAACCCATAGCATCCAGTCTTCTGTACCCAGGT	900				
Qy	922	AGCAGCACTCGAAACTTGGATTTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA	981				
Db	901	AGCAGCACTCGAAACTTGGATTTCAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA	960				
Qy	982	TTTGTCTAAATGACACAGCATACATCTACATGTTTAAATATGACTGGGTGGAAGACTTCT	1041				
Db	961	TTTGTCTAAATGACACAGCATACATCTACATGTTTAAATATGACTGGGTGGAAGACTTCT	1020				

Qy	1042	CCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTATGTAAATCCAGCCTTAGT	1101
Db	1021	CCAGTGGCTATTTTCAATGGACACAGAACTCTACCTTTTATGTAAATCCAGCCTTAGT	1080
Qy	1102	CCAGATGACCAAGTTTATAGTCAGTGGCTCAAGTGATGAAGCTGCCTACATATGGAAGTTC	1161
Db	1081	CCAGATGACCAAGTTTATAGTCAGTGGCTCAAGTGATGAAGCTGCCTACATATGGAAGTTC	1140
Qy	1162	TCCAACCTGGGCAACCTCTCTACTGTCTCTGGGTCAATCTCAAGAGGTCAAGTCTGTG	1221
Db	1141	TCCAACCTGGGCAACCTCTCTACTGTCTCTGGGTCAATCTCAAGAGGTCAAGTCTGTG	1200
Qy	1222	TGCTGTGTCTCATCTGACTTCAAAAGATTTCTACTCTGTCTGTATGACAAATACATAAAA	1281
Db	1201	TGCTGTGTCTCATCTGACTTCAAAAGATTTCTACTCTGTCTGTATGACAAATACATAAAA	1260
Qy	1282	ATCTGGGCTTTGAATAGAGGCTTAGAGGAGAAACAGGAGGTGATAAATTTTCCACGGTG	1341
Db	1261	ATCTGGGCTTTGAATAGAGGCTTAGAGGAGAAACAGGAGGTGATAAATTTTCCACGGTG	1320
Qy	1342	GGTTGGGCTCTCAGAGAAAGAGTCAAGACCTGGCCCTAGTAAACAGTAAACGAGTAGC	1401
Db	1321	GGTTGGGCTCTCAGAGAAAGAGTCAAGACCTGGCCCTAGTAAACAGTAAACGAGTAGC	1380
Qy	1402	CAGAGTACTCTCGCAAGCCCGCAGAGTAAAGTGCAATCCATCCAATTTTCCCGTCA	1461
Db	1381	CAGAGTACTCTCGCAAGCCCGCAGAGTAAAGTGCAATCCATCCAATTTTCCCGTCA	1440
Qy	1462	TCCGAGCTTGTGCCCCCAAGCTGTCTGGAGACCTCCCTCTTCTTCAAAATCTCTACG	1521
Db	1441	TCCGAGCTTGTGCCCCCAAGCTGTCTGGAGACCTCCCTCTTCTTCAAAATCTCTACG	1500
Qy	1522	TTCTCTATTAACACCTCTCTGCGCAAGGCCCGGTCTCCCATCAACAGAGAGGTCTGTCTC	1581
Db	1501	TTCTCTATTAACACCTCTCTGCGCAAGGCCCGGTCTCCCATCAACAGAGAGGTCTGTCTC	1560
Qy	1582	TCCTCCGCTCTCCCAAGCCACCTTCATCTTTCAAGATGTCGATTTAGAAAATGGGTGACC	1641
Db	1561	TCCTCCGCTCTCCCAAGCCACCTTCATCTTTCAAGATGTCGATTTAGAAAATGGGTGACC	1620
Qy	1642	CGAACACCTTCTCTCATCACCACCTCATCTCCACCTGCTCGGAGACCAAGATCATGTCT	1701
Db	1621	CGAACACCTTCTCTCATCACCACCTCATCTCCACCTGCTCGGAGACCAAGATCATGTCT	1680
Qy	1702	CCGAGAAAAGCCCTTATCTCTGTCAGCCAGAGTCAATCCCAAGCAGAGGCTTGTCTCTAG	1761
Db	1681	CCGAGAAAAGCCCTTATCTCTGTCAGCCAGAGTCAATCCCAAGCAGAGGCTTGTCTCTAG	1740
Qy	1762	TCTAGAAATAGATAAGAGAGGAGCTAGACTCAAGCTGTCTGGAGAGTGTGAAAACAAAG	1821
Db	1741	TCTAGAAATAGATAAGAGAGGAGCTAGACTCAAGCTGTCTGGAGAGTGTGAAAACAAAG	1800
Qy	1822	TGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGAATGCCCAAGTTGAAAATCTTTCATTG	1881
Db	1801	TGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGAATGCCCAAGTTGAAAATCTTTCATTG	1860
Qy	1882	GATCTGTGCTGCTCTGTCTGTGTAAACAGGAGCTTGTAGTAGAGTCTCTAGTCTCTTACC	1941
Db	1861	GATCTGTGCTGCTCTGTCTGTGTAAACAGGAGCTTGTAGTAGAGTCTCTAGTCTCTTACC	1920
Qy	1942	AAATCAAGCAAAATTTGAAGGAGCTGGTACAGTATCTCAGAGCCCTCCCTCTATCAGT	2001
Db	1921	AAATCAAGCAAAATTTGAAGGAGCTGGTACAGTATCTCAGAGCCCTCCCTCTATCAGT	1980
Qy	2002	CCGTATGCTTTCAGAAAGCTGTGGAAACGCTACCTCTCTTCTTGGAGACCTTGTGGAGAGG	2061
Db	1981	CCGTATGCTTTCAGAAAGCTGTGGAAACGCTACCTCTCTTCTTGGAGACCTTGTGGAGAGG	2040
Qy	2062	TCTGAAATGTTGAGCAAGAGATAGTTTCCCGAGAGATTAATAACTGTTGTGGCCATG	2121
Db	2041	TCTGAAATGTTGAGCAAGAGATAGTTTCCCGAGAGATTAATAACTGTTGTGGCCATG	2100
Qy	2122	GCAGCAAAACGGAAGGCTGAGAAATCCATCTCCAGAACTCCGTATCCAGACACCCCAAT	2181

Db 2101 GCAGCAAAACGGAAGGCTGAGATCCATCTCCAGGAAGTCGGTCATCCAGACACCCCAAT 2160
Qy 2182 TCAGAGAGACAGAGCGGAAGACATTCGCAAGCCGGTCACATACGCGCCAGCTCCATG 2241
Db 2161 TCAGAGAGACAGAGCGGAAGACATTCGCAAGCCGGTCACATACGCGCCAGCTCCATG 2220
Qy 2242 AGGAAATCTGCACATACCTCCATAGAAAGTCCAGAGGACTTCCTGCTCTGAACAC 2301
Db 2221 AGGAAATCTGCACATACCTCCATAGAAAGTCCAGAGGACTTCCTGCTCTGAACAC 2280
Qy 2302 TCAACAGAAATATAGATTTCTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTAAACAAAG 2361
Db 2281 TCAACAGAAATATAGATTTCTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTAAACAAAG 2340
Qy 2362 CTGAGCTTTGGTCCACTAAACAGAGATGAAATATACAGAGTGACTCTATTAACCTGGTC 2421
Db 2341 CTGAGCTTTGGTCCACTAAACAGAGATGAAATATACAGAGTGACTCTATTAACCTGGTC 2400
Qy 2422 TTTAAGAAAGCTGCCCTTTTCATTTTATAGACAAATCTTTTCAACGCTGAAATGTACTAA 2481
Db 2401 TTTAAGAAAGCTGCCCTTTTCATTTTATAGACAAATCTTTTCAACGCTGAAATGTACTAA 2460
Qy 2482 TCTGGTTCTACTACCAATATGTATATGACGCTTTCCGAGGATGAATGCTGTGTAAAT 2541
Db 2461 TCTGGTTCTACTACCAATATGTATATGACGCTTTCCGAGGATGAATGCTGTGTAAAT 2520
Qy 2542 TCATAAGTAAATTTGTGACTCTAGCATTTTGAATGAATGCTTCACTTTTAAATAT 2601
Db 2521 TCATAAGTAAATTTGTGACTCTAGCATTTTGAATGAATGCTTCACTTTTAAATAT 2580
Qy 2602 TCATCTTCTATAATAATGACATCCAGTTCAATGAGGCAAAACAAAGTTCTTGTTA 2661
Db 2581 TCATCTTCTATAATAATGACATCCAGTTCAATGAGGCAAAACAAAGTTCTTGTTA 2640
Qy 2662 TCTGAAACTTTCTATGCTCAGTGGAAAGTATCTGCAGCCACAGCATGAGGCTGTGAA 2721
Db 2641 TCTGAAACTTTCTATGCTCAGTGGAAAGTATCTGCAGCCACAGCATGAGGCTGTGAA 2700
Qy 2722 GGTGACTGAGAAATCCTCTGCTGAAAGCCCTGGTTCTGTTCTGCTCAACATGATA 2781
Db 2701 GGTGACTGAGAAATCCTCTGCTGAAAGCCCTGGTTCTGTTCTGCTCAACATGATA 2760
Qy 2782 ATTTATTTGAAATACATAATCTTTTCACTATG 2814
Db 2761 ATTTATTTGAAATACATAATCTTTTCACTATG 2793

RESULT 10

US-10-172-118-1645
; Sequence 1645, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linesley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1645
; LENGTH: 4221
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:

; DATABASE ACCESSION NUMBER: NM 016448
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1645

Query Match 98.5%; Score 2789.8; DB 7; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 22 CGATAACGATTTGCTCTGAGAGGCGCAACGTCGATTTCTGCTGAACTTCGAGGCAT 81
Db 1 CGATAACGATTTGCTCTGAGAGGCGCAACGTCGATTTCTGCTGAACTTCGAGGCAT 60
Qy 82 TCTACGACTTTCTCTCAGCTGAGGCTTTTCTCCGACCTGATGCTCTTCAATTCGGTG 141
Db 61 TCTACGACTTTCTCTCAGCTGAGGCTTTTCTCCGACCTGATGCTCTTCAATTCGGTG 120
Qy 142 CTCGCCACGCCACGCTTTGGGCTCTGAGAAATGATGGTCTTCAATACCTCTTCAA 201
Db 121 CTCGCCACGCCACGCTTTGGGCTCTGAGAAATGATGGTCTTCAATACCTCTTCAA 180
Qy 202 TCCCTTCTGACTGGTTATCAGTGACGTGTATGATGAACACACTTCTTATGGAAACA 261
Db 181 TCCCTTCTGACTGGTTATCAGTGACGTGTATGATGAACACACTTCTTATGGAAACA 240
Qy 262 GGAGTCCCAGTTCCTCTTTGGATGTACTTCTCTGCTCCCAATATGAAACATGTA 321
Db 241 GGAGTCCCAGTTCCTCTTTGGATGTACTTCTCTGCTCCCAATATGAAACATGTA 300
Qy 322 CTAGCAGTTGCAATGAAAGGCTTTGTCGATTTGATATAACACAGAAATCAAAAGTTTC 381
Db 301 CTAGCAGTTGCAATGAAAGGCTTTGTCGATTTGATATAACACAGAAATCAAAAGTTTC 360
Qy 382 AGAAAGAGTCTCTCAAGAAATGAGTGCTCACTGGAATCCGCTTTTGACCTGCCCTGG 441
Db 361 AGAAAGAGTCTCTCAAGAAATGAGTGCTCACTGGAATCCGCTTTTGACCTGCCCTGG 420
Qy 442 GTTCTGCTGAACTTTAAACTTTGTACAGCAGCAGGTGATCAAAACAGCAAAATTTTGGGAC 501
Db 421 GTTCTGCTGAACTTTAAACTTTGTACAGCAGCAGGTGATCAAAACAGCAAAATTTTGGGAC 480
Qy 502 GTAAAAGCTGGTGAGCTGATTTGGAAACATGCAAAAGTCAATGAGGCTCAAGTCAGTT 561
Db 481 GTAAAAGCTGGTGAGCTGATTTGGAAACATGCAAAAGTCAATGAGGCTCAAGTCAGTT 540
Qy 562 GCCTTTTCTAAGTTTGAGAAAGCTGATTTCTGTCGGGTGGAGAGATGGCAACATTATG 621
Db 541 GCCTTTTCTAAGTTTGAGAAAGCTGATTTCTGTCGGGTGGAGAGATGGCAACATTATG 600
Qy 622 GTCTGGGATACCAGGTGCAACAAAAAGATGGTGTATAGGCAAGTGAATCAAAATCAGT 681
Db 601 GTCTGGGATACCAGGTGCAACAAAAAGATGGTGTATAGGCAAGTGAATCAAAATCAGT 660
Qy 682 GGAGCTCACAATACCTCAGACAAAGCAAAACCCCTTCAAAACCCCAAGAAACAGAAATCA 741
Db 661 GGAGCTCACAATACCTCAGACAAAGCAAAACCCCTTCAAAACCCCAAGAAACAGAAATCA 720
Qy 742 AAAGGACTTGCTCTCTGTCGATTTCCAGCAAGTGTACTGTGCTCTCTTCAAGAC 801
Db 721 AAAGGACTTGCTCTCTGTCGATTTCCAGCAAGTGTACTGTGCTCTCTTCAAGAC 780
Qy 802 GAGATACCTTAGTCTCAGCAGGAGCTGTGGATGGATAATCAAAAGTATGGGATTTACGT 861
Db 781 GAGATACCTTAGTCTCAGCAGGAGCTGTGGATGGATAATCAAAAGTATGGGATTTACGT 840
Qy 862 AAGAAATTATCTGCTTATCGAACAAGACCCATAGCATCCAAGTCTTTCTGTACCCAGT 921
Db 841 AAGAAATTATCTGCTTATCGAACAAGACCCATAGCATCCAAGTCTTTCTGTACCCAGT 900
Qy 922 AGCAGACTCGAAACTTGGATATTCAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 981
Db 901 AGCAGACTCGAAACTTGGATATTCAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 960
Qy 982 TTTGCTAATTGCACAGACGATAACATCTACATGTTTATATGACTGGGTGGAAGACTTCT 1041

||||| 961 TTTGCTAATTGCACAGCGATTAACATCTACATGTTTATATGACTGGGTGAGACTTCT 1020
||||| 1042 CCAGTGGCTATTTTCAATGAGACACAGAACTCTACTCTTTTATGTAAATCCAGCCTTAGT 1101
||||| 1021 CCAGTGGCTATTTTCAATGAGACACAGAACTCTACTCTTTTATGTAAATCCAGCCTTAGT 1080
||||| 1102 CCAGATGACCACTTTTATGATGAGTCTCAAGTGAAGCTGCCTACATATGGAAGGTC 1161
||||| 1081 CCAGATGACCACTTTTATGATGAGTCTCAAGTGAAGCTGCCTACATATGGAAGGTC 1140
||||| 1162 TCCACACCTGCGAACCTCTACTGTGCTCTGGGTCACTTCTCAAGAGGTCAAGTCTGTG 1221
||||| 1141 TCCACACCTGCGAACCTCTACTGTGCTCTGGGTCACTTCTCAAGAGGTCAAGTCTGTG 1200
||||| 1222 TGCTGGTGTCCATCTGACTTCCAAAGATGTGCTACTCTGTGTTGATGACAATACACTAAAA 1281
||||| 1201 TGCTGGTGTCCATCTGACTTCCAAAGATGTGCTACTCTGTGTTGATGACAATACACTAAAA 1260
||||| 1282 ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGAGAGGTGATAACTTTCCACGGTG 1341
||||| 1261 ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACAGAGAGGTGATAACTTTCCACGGTG 1320
||||| 1342 GGTGGGCTCTCAGAGAGAAAAAGAGTCAAGACCTGGCTAGTAAACAGATACAGAGTAGC 1401
||||| 1321 GGTGGGCTCTCAGAGAGAAAAAGAGTCAAGACCTGGCTAGTAAACAGATACAGAGTAGC 1380
||||| 1402 CAGAGTACTCTGCCAAGCCCCCAGGGTAAAGTGAATCCATCCAAATCTTTCCCGGTCA 1461
||||| 1381 CAGAGTACTCTGCCAAGCCCCCAGGGTAAAGTGAATCCATCCAAATCTTTCCCGGTCA 1440
||||| 1462 TCCGAGCTTGTGCCCAAGCTGTGCTGGAGACCTCCCTCTTCTTCAAATACTCTAGG 1521
||||| 1441 TCCGAGCTTGTGCCCAAGCTGTGCTGGAGACCTCCCTCTTCTTCAAATACTCTAGG 1500
||||| 1522 TTCTCTATTAAACCTCTCTGCCAAGCCCGGTCTCCCATCAACAGAGAGGCTCTGTG 1581
||||| 1501 TTCTCTATTAAACCTCTCTGCCAAGCCCGGTCTCCCATCAACAGAGAGGCTCTGTG 1560
||||| 1582 TCCTCGCTCTCTCCCAAGCACCTTCTCATCTTTCAAGATGTGCGATTAGAACTGGGTGACC 1641
||||| 1561 TCCTCGCTCTCTCCCAAGCACCTTCTCATCTTTCAAGATGTGCGATTAGAACTGGGTGACC 1620
||||| 1642 CGAACACCTTCTCATCACCACCTCATCTCACTCACTGCTTGGAGACCGAAGATCATGTCT 1701
||||| 1621 CGAACACCTTCTCATCACCACCTCATCTCACTCACTGCTTGGAGACCGAAGATCATGTCT 1680
||||| 1702 CCGAGAAAGCCCTTATTCTGTGAGCCAGAGTGTATCCCAAGCAGAGGCTTGTCTGTAG 1761
||||| 1681 CCGAGAAAGCCCTTATTCTGTGAGCCAGAGTGTATCCCAAGCAGAGGCTTGTCTGTAG 1740
||||| 1762 TCTAGAAATAGAGTAAAGAGGCTTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAG 1821
||||| 1741 TCTAGAAATAGAGTAAAGAGGCTTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAG 1800
||||| 1822 TGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGTATGGCCAAAGTTGAAATCTTCAATTG 1881
||||| 1801 TGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGTATGGCCAAAGTTGAAATCTTCAATTG 1860
||||| 1882 GATCTGTGCTGCTTGTCTGTAACCAAGGAAGACCTTGTATGAGGACTCTCTAGGCTCTTACC 1941
||||| 1861 GATCTGTGCTGCTTGTCTGTAACCAAGGAAGACCTTGTATGAGGACTCTCTAGGCTCTTACC 1920
||||| 1942 AAATCAAGCAAAATTTGAAGAGCTGGTACAGTATCTCAGAGCCCTCCGCTCTCTATCAGT 2001
||||| 1921 AAATCAAGCAAAATTTGAAGAGCTGGTACAGTATCTCAGAGCCCTCCGCTCTCTATCAGT 1980
||||| 2002 CCGTATGCTTTCAGAAAGCTGTGAACGCTTACCTCTTCTTTCAGACCTTGTGGAGAGGG 2061
||||| 1981 CCGTATGCTTTCAGAAAGCTGTGAACGCTTACCTCTTCTTTCAGACCTTGTGGAGAGGG 2040
||||| 2062 TCTGAAATGGTAGGCAAGAGAAATAGTTCCTCCAGAGAAATAAAACTGGTGTGTCGCAATG 2121
|||||

Db 2041 TCTGAAATGGTAGGCAAGAGAAATAGTTCCTCCAGAGAAATAAAAACTGGTGTGTCGCCATG 2100
Qy 2122 GCAGCCAAAACGGAAGGCTGAGAAATCCATCTCCAGAAAGTCCGTCTATCCACAGACCCCAAT 2181
Db 2101 GCAGCCAAAACGGAAGGCTGAGAAATCCATCTCCAGAAAGTCCGTCTATCCACAGACCCCAAT 2160
Qy 2182 TCCAGGAGACAGAGCGGAAGAGACATTCGCAAGCCCGGTCCACATCAGGCCCGAGCTCCATG 2241
Db 2161 TCCAGGAGACAGAGCGGAAGAGACATTCGCAAGCCCGGTCCACATCAGGCCCGAGCTCCATG 2220
Qy 2242 AGGAAATCTCCACATACTCTCCATAGAAAGTCCAGGAGGACTTCTGTGGTCCCTGAAACAC 2301
Db 2221 AGGAAATCTCCACATACTCTCCATAGAAAGTCCAGGAGGACTTCTGTGGTCCCTGAAACAC 2280
Qy 2302 TCAACAGAAATATAGATCTTAATCTGAGTGAATTAATCTGAGCTTTGGTCCACTAAACAAAG 2361
Db 2281 TCAACAGAAATATAGATCTTAATCTGAGTGAATTAATCTGAGCTTTGGTCCACTAAACAAAG 2340
Qy 2362 CTGAGCTTTGGTCCACTAAACAGATCAAAATACAAAGAGTGAATCTATAACTCTGTGTC 2421
Db 2341 CTGAGCTTTGGTCCACTAAACAGATCAAAATACAAAGAGTGAATCTATAACTCTGTGTC 2400
Qy 2422 TTTAAGAAAGCTGCGCTTTTCAATTTTATAGACAAAATCTTTTCAACGCTGAAATGTACTAA 2481
Db 2401 TTTAAGAAAGCTGCGCTTTTCAATTTTATAGACAAAATCTTTTCAACGCTGAAATGTACTAA 2460
Qy 2482 TCTGGTCTCTACATCAATATGATATGAGCTTCCGAGGATGAATGCTGTGTTAAATTT 2541
Db 2461 TCTGGTCTCTACATCAATATGATATGAGCTTCCGAGGATGAATGCTGTGTTAAATTT 2520
Qy 2542 TCATAAAGTAAATTTGTACACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTTAT 2601
Db 2521 TCATAAAGTAAATTTGTACACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTTAT 2580
Qy 2602 TCATCTCTCTATATATATGACATCCAGTTCATGGAGGCAAAAACAAAGTTTCTTGTGA 2661
Db 2581 TCATCTCTCTATATATATGACATCCAGTTCATGGAGGCAAAAACAAAGTTTCTTGTGA 2640
Qy 2662 TCCTGAAACTTTCTATGCTGAGTGAAGATCTGCGAGGCCACAGCATGAGGCTGTGAA 2721
Db 2641 TCCTGAAACTTTCTATGCTGAGTGAAGATCTGCGAGGCCACAGCATGAGGCTGTGAA 2700
Qy 2722 GGCTGACTGAGAAATCCCTGCTGAGACCCCTGCTTCTGCTCCCAACATGTATA 2781
Db 2701 GGCTGACTGAGAAATCCCTGCTGAGACCCCTGCTTCTGCTCCCAACATGTATA 2760
Qy 2782 ATTTATTTGAAATACATAATCTTTTCACTATG 2814
Db 2761 ATTTATTTGAAATACATAATCTTTTCACTATG 2793

RESULT 11

US-10-342-887-1645
; Sequence 1645, Application US/10342387
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Robert, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118

;; PRIOR FILING DATE: 2002-06-14
;; NUMBER OF SEQ ID NOS: 2699
;; SEQ ID NO 1645
;; LENGTH: 4221
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-342-887-1645

Query Match 98.5%; Score 2789, 8; DB 8; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	22	CGATAACGATTTGTGTGTCAGAGGCGCAACGTCGCAATTTCTGCTGAACCTTGGAGCAAT	81
Db	1	CGATAACGATTTGTGTGTCAGAGGCGCAACGTCGCAATTTCTGCTGAACCTTGGAGCAAT	60
Qy	82	TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCTGATGCTCTTCAATTCGGTG	141
Db	61	TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCTGATGCTCTTCAATTCGGTG	120
Qy	142	CTCGCCAGCCCGCAGCTTGGCGTCTCGAGAAATGGATGGTCTTCAACAATACCCCTCTTCAA	201
Db	121	CTCGCCAGCCCGCAGCTTGGCGTCTCGAGAAATGGATGGTCTTCAACAATACCCCTCTTCAA	180
Qy	202	TCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACAACACTTCTTATGAGAAACA	261
Db	181	TCCTTCTGACTGGTTATCAGTGCAGTGGTAAATGATGAACAACACTTCTTATGAGAAACA	240
Qy	262	GGAGTCCAGTTCCCTCTTTTGGATGTAACCTTCTCTTCTGCTCCCAATATGGAACATGTA	321
Db	241	GGAGTCCAGTTCCCTCTTTTGGATGTAACCTTCTCTTCTGCTCCCAATATGGAACATGTA	300
Qy	322	CTAGCAGTTGCCAATGAAGAAGCTTGTTCGATGTAATGAACACAGATCAACAAGTTTC	381
Db	301	CTAGCAGTTGCCAATGAAGAAGCTTGTTCGATGTAATGAACACAGATCAACAAGTTTC	360
Qy	382	AGAAAGAAGTGTCTCAAGAATGGATGGCTCACTGGAAATGCCGTCTTTGACCTGGCGCTGG	441
Db	361	AGAAAGAAGTGTCTCAAGAATGGATGGCTCACTGGAAATGCCGTCTTTGACCTGGCGCTGG	420
Qy	442	GTTCTCGGTGAACCTTAAACTTTGTTACAGCAGCAGGTGATCAAAACAGCAAAATTTGGGAC	501
Db	421	GTTCTCGGTGAACCTTAAACTTTGTTACAGCAGCAGGTGATCAAAACAGCAAAATTTGGGAC	480
Qy	502	GTAAGAAGCTGTCAGCTGTAAGCAATGCAAGAGTCAATCAATGACGCTCAAGTCAGTT	561
Db	481	GTAAGAAGCTGTCAGCTGTAAGCAATGCAAGAGTCAATCAATGACGCTCAAGTCAGTT	540
Qy	562	GCCTTTTCTAAGTTTGAGAAAGCTGTATTCTGTACGGGTGGAAGAGATGGCAACATTATG	621
Db	541	GCCTTTTCTAAGTTTGAGAAAGCTGTATTCTGTACGGGTGGAAGAGATGGCAACATTATG	600
Qy	622	GTCCTGGGATACCAAGTGCACAAAAAGATGGGTTTTATAGGCAAGTGAATCAAAATCAGT	681
Db	601	GTCCTGGGATACCAAGTGCACAAAAAGATGGGTTTTATAGGCAAGTGAATCAAAATCAGT	660
Qy	682	GGAGCTCACATACCTTCAGACAGCAACCCCTTCAAAACCCCAAGAAACAGAAATTC	741
Db	661	GGAGCTCACATACCTTCAGACAGCAACCCCTTCAAAACCCCAAGAAACAGAAATTC	720
Qy	742	AAAGGACTTGTCTCTCTGTGGATTTCCAGCAAGTGTACTGTGCTCTTTTCAAGAC	801
Db	721	AAAGGACTTGTCTCTCTGTGGATTTCCAGCAAGTGTACTGTGCTCTTTTCAAGAC	780
Qy	802	GAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAAGTATGGGATTTACGT	861
Db	781	GAGAAATACCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAAGTATGGGATTTACGT	840
Qy	862	AAGAATATACCTTATCGACAAGAACCCATAGCATCCAGTCTTCTGTACCCAGGT	921
Db	841	AAGAATATACCTTATCGACAAGAACCCATAGCATCCAGTCTTCTGTACCCAGGT	900
Qy	922	AGCAGCACTCGAAAACTTGGATATTCAAGTCTGATTTTGGATTCCACTGGCTCTACTTTA	981

Db	901	AGCAGCACTCGAAAACTTGGATATTCAAGTCTGATTTTGGATTCCACTGGCTCTACTTTA	960
Qy	982	TTTGTCTAATTTGCACAGACGATTAACATCTACATCTGTTTAAATATGACCTGGGTTGAAGCTTCT	1041
Db	961	TTTGTCTAATTTGCACAGACGATTAACATCTACATCTGTTTAAATATGACCTGGGTTGAAGCTTCT	1020
Qy	1042	CCAGTGGCTATTCTCAATGGACACCAAGAACTCTACCTTTTATGTAAATTCACAGCCTTAGT	1101
Db	1021	CCAGTGGCTATTCTCAATGGACACCAAGAACTCTACCTTTTATGTAAATTCACAGCCTTAGT	1080
Qy	1102	CCAGATGACCCAGTTTTTAGTCAAGTGGCTCAAGTGAATGAAGCTGCCCTACATATGGAAGGTC	1161
Db	1081	CCAGATGACCCAGTTTTTAGTCAAGTGGCTCAAGTGAATGAAGCTGCCCTACATATGGAAGGTC	1140
Qy	1162	TCACACCCCTGGCAACCTCTACTGCTCTCTGGGTCATCTCAAGAGGTCAGCTCTGTG	1221
Db	1141	TCACACCCCTGGCAACCTCTACTGCTCTCTGGGTCATCTCAAGAGGTCAGCTCTGTG	1200
Qy	1222	TGCTGGTGTCCATCTGACTTCAAAAGATTGCTACTCTGTCTGTGATGACAAATACACTAAAA	1281
Db	1201	TGCTGGTGTCCATCTGACTTCAAAAGATTGCTACTCTGTCTGTGATGACAAATACACTAAAA	1260
Qy	1282	ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACCAAGGAGGTGATAAACTTTCCAGGTC	1341
Db	1261	ATCTGGCGCTTGAATAGAGGCTTAGAGGAGAAACCAAGGAGGTGATAAACTTTCCAGGTC	1320
Qy	1342	GGTTGGGCTCTCAGAGAAGAAAAGAGTCAAGACCTGGCCTAGTAAACAGTACGAGTAGC	1401
Db	1321	GGTTGGGCTCTCAGAGAAGAAAAGAGTCAAGACCTGGCCTAGTAAACAGTACGAGTAGC	1380
Qy	1402	CAGAGTACTCTCGCAAGCCCCCAGGGTAAAGTGAACATCAATCAAACTTTCCCGCTCA	1461
Db	1381	CAGAGTACTCTCGCAAGCCCCCAGGGTAAAGTGAACATCAATCAAACTTTCCCGCTCA	1440
Qy	1462	TCCGACAGCTTGTGCCCAAGCTGTCTGGAGACCTCCCTCTTCTTCAAAATACTCCTACG	1521
Db	1441	TCCGACAGCTTGTGCCCAAGCTGTCTGGAGACCTCCCTCTTCTTCAAAATACTCCTACG	1500
Qy	1522	TTCTCTATTAAACCTCTCTGCGCAAGCCCGGCTCTCCCATCAACAGAGAGGCTCTGTG	1581
Db	1501	TTCTCTATTAAACCTCTCTGCGCAAGCCCGGCTCTCCCATCAACAGAGAGGCTCTGTG	1560
Qy	1582	TCCTCGCTCTCCCAAGCCACCTTCACTTTCAAGATGTGCTATTAGAAAATGGGGTAC	1641
Db	1561	TCCTCGCTCTCCCAAGCCACCTTCACTTTCAAGATGTGCTATTAGAAAATGGGGTAC	1620
Qy	1642	CGAACACCTTCTCTATCAACCCATCACTCCACCTGTCTCGGAGACCAAGATCATGTCT	1701
Db	1621	CGAACACCTTCTCTATCAACCCATCACTCCACCTGTCTCGGAGACCAAGATCATGTCT	1680
Qy	1702	CCGAGAAAAGCCCTTATTTCTGTGAGCCAGAAAGTCAATCCCAAGAGAGGCTTGTCTGAG	1761
Db	1681	CCGAGAAAAGCCCTTATTTCTGTGAGCCAGAAAGTCAATCCCAAGAGAGGCTTGTCTGAG	1740
Qy	1762	TCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAG	1821
Db	1741	TCTAGAAATAGAGTAAAGAGGAGGCTAGACTCAAGCTGTCTGGAGAGTGTGAAACAAAG	1800
Qy	1822	TGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGTATGGCCCAAGTTGAAAATCTTCAATTTG	1881
Db	1801	TGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGTATGGCCCAAGTTGAAAATCTTCAATTTG	1860
Qy	1882	GATCTGTGCTGTGCTGGTAAACAGGAAGACCTTTAGTAGGACTCTCTAGGCTCTTACC	1941
Db	1861	GATCTGTGCTGTGCTGGTAAACAGGAAGACCTTTAGTAGGACTCTCTAGGCTCTTACC	1920
Qy	1942	AAATCAAGCAAAATTTGAAGGAGCTGTACCAAGTATCTCAGAGCCTCCGCTCTCTATCAGT	2001
Db	1921	AAATCAAGCAAAATTTGAAGGAGCTGTACCAAGTATCTCAGAGCCTCCGCTCTCTATCAGT	1980
Qy	2002	CCGTATGCTTCAGAAAAGCTGTGGAACGCTACTCTTCTCTTTGAGACCTTGTGGAGAGGG	2061

1981 CCGTATGCTTCAGAAAGCTGTGGAAGCTACCTCTCTCTTTGAGACCTTGTGGAGAGGG 2040
2062 TCTGAAATGGTAGGCAAGAGAAATAGTTCCTCCAGAGAAATAAAAACTGGTGTGGCCATG 2121
2041 TCTGAAATGGTAGGCAAGAGAAATAGTTCCTCCAGAGAAATAAAAACTGGTGTGGCCATG 2100
2122 GCAGCCAAACGGAAGCTGAGAAATCCATCTCCAGAAAGTCCGTCATCCAGACACCCCAAT 2181
2101 GCAGCCAAACGGAAGCTGAGAAATCCATCTCCAGAAAGTCCGTCATCCAGACACCCCAAT 2160
2182 TCAGGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTCAACCATCAAGCCGCTCAGTCCCATG 2241
2161 TCAGGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTCAACCATCAAGCCGCTCAGTCCCATG 2220
2242 AGGAAATCTGCACATACCTTCATAGAAAGTCCAGAGGAGATCTTGTTGGTCTCTGAAACAC 2301
2221 AGGAAATCTGCACATACCTTCATAGAAAGTCCAGAGGAGATCTTGTTGGTCTCTGAAACAC 2280
2302 TCAACAGAAATATAGAAATCTTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTAAACCAAG 2361
2281 TCAACAGAAATATAGAAATCTTAATCTGAGTGAGTTACTGAGCTTTGGTCCACTAAACCAAG 2340
2362 CTGAGCTTTGGTCCACTAAACCAAGATGAAATACAAAGTGAATCTATAACTCTGGTTC 2421
2341 CTGAGCTTTGGTCCACTAAACCAAGATGAAATACAAAGTGAATCTATAACTCTGGTTC 2400
2422 TTTAAGAAAGCTGCCCTTTTCAATTTTATAGACAAAATCTTTTCAACGCTGAAATGTACCTAA 2481
2401 TTTAAGAAAGCTGCCCTTTTCAATTTTATAGACAAAATCTTTTCAACGCTGAAATGTACCTAA 2460
2482 TCTGGTCTCTACCAATATATATATGAGCTTCCGAGAGATGAAATGCTGTTTAAATTT 2541
2461 TCTGGTCTCTACCAATATATATATGAGCTTCCGAGAGATGAAATGCTGTTTAAATTT 2520
2542 TCATAAGTAAATTTCTGCTACTCTAGCAATTTTGAATGAATAGTCTTCACTTTTAAATAT 2601
2521 TCATAAGTAAATTTCTGCTACTCTAGCAATTTTGAATGAATAGTCTTCACTTTTAAATAT 2580
2602 TCATCTTCTCTAATAATGACATCCAGTTCATGAGGCAAAAAACAAGTTTCTTGTTA 2661
2581 TCATCTTCTCTAATAATGACATCCAGTTCATGAGGCAAAAAACAAGTTTCTTGTTA 2640
2662 TCCGGAACCTTTCTATGCTCAGTGGAAGTATCTGCCAGCCACAGCATGAGCCCTGTGAA 2721
2641 TCCGGAACCTTTCTATGCTCAGTGGAAGTATCTGCCAGCCACAGCATGAGCCCTGTGAA 2700
2722 GGCTGACTGAGAAATCCTCTGCTGGAAGACCCCTGGTTCGTGCTCCACATGTATA 2781
2701 GGCTGACTGAGAAATCCTCTGCTGGAAGACCCCTGGTTCGTGCTCCACATGTATA 2760
2782 ATTTTATTTGAAATACATAATCTTTTCACTATG 2814
2761 ATTTTATTTGAAATACATAATCTTTTCACTATG 2793

RESULT 12
US-10-848-755A-179
; Sequence 179, Application US/10848755A
; Publication No. US20050054826A1
; GENERAL INFORMATION:
; APPLICANT: Mao, Mao
; TITLE OF INVENTION: HUMAN DIAPHANOUS-3 GENE AND METHODS OF USE THEREFOR
; FILE REFERENCE: 9301-196-999
; CURRENT APPLICATION NUMBER: US/10/848, 755A
; CURRENT FILING DATE: 2004-05-18
; PRIOR APPLICATION NUMBER: 60/471,842
; PRIOR FILING DATE: 2003-05-11
; NUMBER OF SEQ ID NOS: 275
; SOFTWARE: PatentIn version 3.2 CAM: 301891-999188
; SEQ ID NO 179
; TYPE: DNA
; ORGANISM: Homo sapiens

US-10-848-755A-179
Query Match 98.5%; Score 2789.8; DB 10; Length 4221;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2791; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 22 CGATAACGATTTGTGTGAGAGCGCAAGCTGCGATTTCTGCTGGAACCTTTGGAGGCAT 81
DB 1 CGATAACGATTTGTGTGAGAGCGCAAGCTGCGATTTCTGCTGGAACCTTTGGAGGCAT 60
QY 82 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG 141
DB 61 TCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGCTCTTCAATTCGGTG 120
QY 142 CTCGCGCAGCCCGAGCTTTGGCGTCTTGAGAAATGAGTGGTCTTACATAACCTCTTCAA 201
DB 121 CTCGCGCAGCCCGAGCTTTGGCGTCTTGAGAAATGAGTGGTCTTACATAACCTCTTCAA 180
QY 202 TCCCTTCTGACTGGTTCAGTGCAGTGGTAAATGATGAACACACTTCTTATGGAGAAACA 261
DB 181 TCCCTTCTGACTGGTTCAGTGCAGTGGTAAATGATGAACACACTTCTTATGGAGAAACA 240
QY 262 GGAGTCCAGTTCCTCTCTTTTGGATGATACCTTCTCTCTGCTCCCAATATGGAACATGTA 321
DB 241 GGAGTCCAGTTCCTCTCTTTTGGATGATACCTTCTCTCTGCTCCCAATATGGAACATGTA 300
QY 322 CTAGCAGTTGCCAATGAAGAAGGCTTTGTTCGATTTGATATACAGAGATCAAAAGTTTC 381
DB 301 CTAGCAGTTGCCAATGAAGAAGGCTTTGTTCGATTTGATATACAGAGATCAAAAGTTTC 360
QY 382 AGAAGAAGTCTCTCAAGAAGATGATGGCTCACTGGAATGCCCTTTTGACTCGGCCTGG 441
DB 361 AGAAGAAGTCTCTCAAGAAGATGATGGCTCACTGGAATGCCCTTTTGACTCGGCCTGG 420
QY 442 GTTCTGCTGAACTTAAACTTTGTACAGCAGCAGTGATCAAAACAGCCAAATTTTGGAC 501
DB 421 GTTCTGCTGAACTTAAACTTTGTACAGCAGCAGTGATCAAAACAGCCAAATTTTGGAC 480
QY 502 GTAAAGCTGGTGTGAGCTGATTTGGAACATGCAAGGTCAATCAATGAGCCTCAAGTCAGT 561
DB 481 GTAAAGCTGGTGTGAGCTGATTTGGAACATGCAAGGTCAATCAATGAGCCTCAAGTCAGT 540
QY 562 GCCTTTTCTAAAGTTTGAGAAAGCTGTATCTGTCGGGTGGAAGAGATGGCAACATTATG 621
DB 541 GCCTTTTCTAAAGTTTGAGAAAGCTGTATCTGTCGGGTGGAAGAGATGGCAACATTATG 600
QY 622 GTCTGGGATACAGGTGCACAAACAAAGATGGTTTTATAGGCAAGTGAATCAAAATCAGT 681
DB 601 GTCTGGGATACAGGTGCACAAACAAAGATGGTTTTATAGGCAAGTGAATCAAAATCAGT 660
QY 682 GGAGCTCAATACCTCTGAGACAGCAACCCCTTCAAAACCCAGAGAAACAGAAATTCATCA 741
DB 661 GGAGCTCAATACCTCTGAGACAGCAACCCCTTCAAAACCCAGAGAAACAGAAATTCATCA 720
QY 742 AAAGGACTTGTCTCTGCTGATTTCCAGCAAAAGTGTACTGTGCTCTCTTTCAAGAC 801
DB 721 AAAGGACTTGTCTCTGCTGATTTCCAGCAAAAGTGTACTGTGCTCTCTTTCAAGAC 780
QY 802 GAGAAATCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAAGTATGGGATTTACGT 861
DB 781 GAGAAATCTTAGTCTCAGCAGGAGCTGTGGATGGGATAATCAAAAGTATGGGATTTACGT 840
QY 862 AAGAAATTAATCTGCTTATCGACAAAGCCCATAGCATCAAGTCTTTTCTGTACCCAGGT 921
DB 841 AAGAAATTAATCTGCTTATCGACAAAGCCCATAGCATCAAGTCTTTTCTGTACCCAGGT 900
QY 922 AGCAGCACTCGAAAACTTTGGATATTAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 981
DB 901 AGCAGCACTCGAAAACTTTGGATATTAAGTCTGATTTTGGATTTCCACTGGCTCTACTTTA 960
QY 982 TTTTGTAAATTCACAGACGATAAACHATCTACATGTTTAAATATGACGTGGGTGAAGCTTCT 1041
DB 961 TTTTGTAAATTCACAGACGATAAACHATCTACATGTTTAAATATGACGTGGGTGAAGCTTCT 1020

```
QY 1042 CCAGTGGCTATTTTCAATGGACACAGAACTTACCTTTTATGTGMAAATCCAGCCTTAGT 1101
Db 1021 CCAGTGGCTATTTTCAATGGACACAGAACTTACCTTTTATGTGMAAATCCAGCCTTAGT 1080
QY 1102 CCAGATGACCAAGTTTTTATGAGTGGCTCAAGTGAAGAGCTGCCTACATATGGAAGGTC 1161
Db 1081 CCAGATGACCAAGTTTTTATGAGTGGCTCAAGTGAAGAGCTGCCTACATATGGAAGGTC 1140
QY 1162 TCCACACCTTGGCAACCTCTTACTGTGCTCTCTGGGTCAATTTCTCAAGAGGTCAAGTCTGTG 1221
Db 1141 TCCACACCTTGGCAACCTCTTACTGTGCTCTCTGGGTCAATTTCTCAAGAGGTCAAGTCTGTG 1200
QY 1222 TGCTGGGTGCTCAATCTGACTTTCAAAAGATTGCTACCTGTTCTGATGACAATACACTAAAA 1281
Db 1201 TGCTGGGTGCTCAATCTGACTTTCAAAAGATTGCTACCTGTTCTGATGACAATACACTAAAA 1260
QY 1282 ATCTGGCGCTTGAATAGAGCTTTAGAGGAGAAACCAAGAGGTTGATAACTTTTCCAGGGTG 1341
Db 1261 ATCTGGCGCTTGAATAGAGCTTTAGAGGAGAAACCAAGAGGTTGATAACTTTTCCAGGGTG 1320
QY 1342 GGTGGGCTTCTCAGAGAAAAGAGTCAAGACCTGGCTAGTACAGTAACAGTAGTC 1401
Db 1321 GGTGGGCTTCTCAGAGAAAAGAGTCAAGACCTGGCTAGTACAGTAACAGTAGTAGC 1380
QY 1402 CAGAGTACTCTCGCAAGCCCTCAGGGTAAAGTGCAATCCATCCAAATTTCTCCCGGTCA 1461
Db 1381 CAGAGTACTCTCGCAAGCCCTCAGGGTAAAGTGCAATCCATCCAAATTTCTCCCGGTCA 1440
QY 1462 TCCGAGCTTGTGCCCAAGCTGTGCTGGAGACCTCCCTCTTCTTCAAAATCTCTCTACG 1521
Db 1441 TCCGAGCTTGTGCCCAAGCTGTGCTGGAGACCTCCCTCTTCTTCAAAATCTCTCTACG 1500
QY 1522 TTCTCTATTAAGCTTCTCTGCAAGCCCGGTCTCCCATCAACAGAGAGCTCTGTC 1581
Db 1501 TTCTCTATTAAGCTTCTCTGCAAGCCCGGTCTCCCATCAACAGAGAGCTCTGTC 1560
QY 1582 TCCTCGTCTCTCCCAAGCCACCTTCATCTTCAAGATGTCGATTAGAAATCGGTGACC 1641
Db 1561 TCCTCGTCTCTCCCAAGCCACCTTCATCTTCAAGATGTCGATTAGAAATCGGTGACC 1620
QY 1642 CGAACACCTTCTCATCACCAACCCATCACTCCACCTGCTTCCGAGACCAAGATCATGTCT 1701
Db 1621 CGAACACCTTCTCATCACCAACCCATCACTCCACCTGCTTCCGAGACCAAGATCATGTCT 1680
QY 1702 CCAGAAAAGCCCTTATCTCTGTGAGCCAGAAGTCATCCCAAGAGAGGCTCTGCTGAG 1761
Db 1681 CCAGAAAAGCCCTTATCTCTGTGAGCCAGAAGTCATCCCAAGAGAGGCTTCTGCTGAG 1740
QY 1762 TCTAGAAATAGAGTAAAGAGGAGCTAGACTCAAGCTGCTGGAGAGTGTGAAACAAAAG 1821
Db 1741 TCTAGAAATAGAGTAAAGAGGAGCTAGACTCAAGCTGCTGGAGAGTGTGAAACAAAAG 1800
QY 1822 TGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGTGAGCCCAAGTTGAAAATCTTCAATTTG 1881
Db 1801 TGTGTGAAGAGTTGTAACTGTGTGACTGAGCTTGTGAGCCCAAGTTGAAAATCTTCAATTTG 1860
QY 1882 GATCTGTGCTGCTTGTGTGTAACAGGAAAGACTTGTAGTAAGAGCTCTCTAGTCTCTTACC 1941
Db 1861 GATCTGTGCTGCTTGTGTGTAACAGGAAAGACTTGTAGTAAGAGCTCTCTAGTCTCTTACC 1920
QY 1942 AAATCAAGCAAAATTGAAGAGCTGTGTACCAAGTATCTCAGAGCTCGTCTCTCTATCAGT 2001
Db 1921 AAATCAAGCAAAATTGAAGAGCTGTGTACCAAGTATCTCAGAGCTCGTCTCTCTATCAGT 1980
QY 2002 CCGTATCTCTTCAAGAACTGTGAAACGCTACCTCTTCTTGTGAGACCTTGTGGAGAAGGG 2061
Db 1981 CCGTATCTCTTCAAGAACTGTGAAACGCTACCTCTTCTTGTGAGACCTTGTGGAGAAGGG 2040
QY 2062 TCTGAAATGTGAGCAAGAGAAATAGTTTCCCAAGAGAAATAAAATCTGTTGTGGCCATG 2121
Db 2041 TCTGAAATGTGAGCAAGAGAAATAGTTTCCCAAGAGAAATAAAATCTGTTGTGGCCATG 2100
```

```
QY 2122 GCAGCCAAAACGGAAGGCTGAGAATCCATCTCCACGAAGTCCGTATCCAGACACACCCAAT 2181
Db 2101 GCAGCCAAAACGGAAGGCTGAGAATCCATCTCCACGAAGTCCGTATCCAGACACACCCAAT 2160
QY 2182 TCAGAGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTCCACATCAGCCCGAGCTCCATG 2241
Db 2161 TCAGAGAGACAGAGCGGAAAGACATTTGCCAAGCCCGGTCCACATCAGCCCGAGCTCCATG 2220
QY 2242 AGGAAATCTGCACATACATTTCCATAGAAAGTCCAGGAGGACTTCTGTGCTCTGAAAC 2301
Db 2221 AGGAAATCTGCACATACATTTCCATAGAAAGTCCAGGAGGACTTCTGTGCTCTGAAAC 2280
QY 2302 TCAACAGAAATTAAGATTTCTAATCTGAGTGAGTTACTGAGCTTTGGTCCACATAAAACAAG 2361
Db 2281 TCAACAGAAATTAAGATTTCTAATCTGAGTGAGTTACTGAGCTTTGGTCCACATAAAACAAG 2340
QY 2362 CTGAGCTTTGGTCCACTAAACCAAGATGAAAATACAGAGTGACTCTATTAATCTGGTTC 2421
Db 2341 CTGAGCTTTGGTCCACTAAACCAAGATGAAAATACAGAGTGACTCTATTAATCTGGTTC 2400
QY 2422 TTTAAGAAAGCTGCCCTTTTCTATTTTGTAGACAAAATCTTTTCAACGCTGAAATGTACTTAA 2481
Db 2401 TTTAAGAAAGCTGCCCTTTTCTATTTTGTAGACAAAATCTTTTCAACGCTGAAATGTACTTAA 2460
QY 2482 TCTGGTTCTACTACCATTAATGTATATGACAGCTTCCCGAGGATGAATGTGTGTTTAAAT 2541
Db 2461 TCTGGTTCTACTACCATTAATGTATATGACAGCTTCCCGAGGATGAATGTGTGTTTAAAT 2520
QY 2542 TCATAAGTAAATTTGTCTACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTAAT 2601
Db 2521 TCATAAGTAAATTTGTCTACTCTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTAAT 2580
QY 2602 TCATCTTCTCTATAATATGACATCCAGTTTCATGAGGCAAAAACAAGTTTCTTGTGTA 2661
Db 2581 TCATCTTCTCTATAATATGACATCCAGTTTCATGAGGCAAAAACAAGTTTCTTGTGTA 2640
QY 2662 TCTGAAACTTTCTATGCTCAGTGGAAAGATATCTGCCAGCCACAGCATGAGGCTGTGAA 2721
Db 2641 TCTGAAACTTTCTATGCTCAGTGGAAAGATATCTGCCAGCCACAGCATGAGGCTGTGAA 2700
QY 2722 GGCTGACTGAGAAATCTCTGCTGAGAGCCCTGGTTCTGTTCTGCTTCCACATGTATA 2781
Db 2701 GGCTGACTGAGAAATCTCTGCTGAGAGCCCTGGTTCTGTTCTGCTTCCACATGTATA 2760
QY 2782 ATTTTATTTGAATACATAATCTTTTCACTATG 2814
Db 2761 ATTTTATTTGAATACATAATCTTTTCACTATG 2793
```

RESULT 13

```
US-10-330-773-673
; Sequence 673, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330,773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 673
; LENGTH: 4083
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-330-773-673
```

Query Match 86.3%; Score 2442.4; DB 11; Length 4083;
Best Local Similarity 93.5%; Pred. No. 0;
Matches 2630; Conservative 0; Mismatches 11; Indels 171; Gaps 1;

QY 3 CACGAGCGGAGTTGGAGGCGGATACGATTGTGTGTGAGAGCGGCAACGTCGCGATTC 62

[illegible]

946	Db	TGTAATAATCCAGCCTTAGTCCAGATGACCACTGTTTTTTAGTCAGTGGCTCAAGTGATGAAGC	1005
1143	Qy	TGCTCATATGAAAGGTCTCCACACCGTGGCAACCTCTACTGTGCTCTCTGGGTCAATTC	1202
1006	Db	TGCCTCATATGGAGAGGTCTCCACACCGTGGCAACCTCTACTGTGCTCTGGGTCAATTC	1065
1203	Qy	TCAAGAGGTCAAGTCTGTGTGTGGTGTCCATCTGACTTCCACAAAGATTGCTACCTGTTC	1262
1066	Db	TCAAGAGGTCAAGTCTGTGTGTGGTGTCCATCTGACTTCCACAAAGATTGCTACCTGTTC	1125
1263	Qy	TGATGACAATACACTAAATAATCTGGCGCTTGAATAGAGGCTTAGAGAGAGAAACGAGGAGG	1322
1126	Db	TGATGACAATACACTAAATAATCTGGCGCTTGAATAGAGGCTTAGAGAGAGAAACGAGGAGG	1185
1323	Qy	TGATAAACTTTTCCACGGTGGGTTGGGCTCTCAGAGAGAAAAAGAGTCAAGACCTGGCCCT	1382
1186	Db	TGATAAACTTTTCCACGGTGGGTTGGGCTCTCAGAGAGAAAAAGAGTCAAGACCTGGCCCT	1245
1383	Qy	AGTAACAGTAGTAAACGAGTAGCCAGAGTACTCTCTGCCAAAGCCCCCAGGGTAAGTGCATTC	1442
1246	Db	AGTAACAGTAGTAAACGAGTAGCCAGAGTACTCTCTGCCAAAGCCCCCAGGGTAAGTGCATTC	1305
1443	Qy	ATCCAAATCTTCCCGTCAATCCGAGCTTGTGGCCCAAGCTGTCTGGAGACCTCCCTCT	1502
1306	Db	ATCCAAATCTTCCCGTCAATCCGAGCTTGTGGCCCAAGCTGTCTGGAGACCTCCCTCT	1365
1503	Qy	TCCTTCAAAATCTCTACGTTCTCTATTAATAAAGCTCTCTGCCAAAGCCCCCGGTCTCCCAT	1562
1366	Db	TCCTTCAAAATCTCTACGTTCTCTATTAATAAAGCTCTCTGCCAAAGCCCCCGGTCTCCCAT	1425
1563	Qy	CAACAGAGAGGCTCTGTCTCTCCGCTCTCTCCCAAGCCACTTTCATCTTTCAAGATGTC	1622
1426	Db	CAACAGAGAGGCTCTGTCTCTCCGCTCTCTCCCAAGCCACTTTCATCTTTCAAGATGTC	1485
1623	Qy	GATTAGAAAATGGGTGACCCGAACACTTCTCTCATACACACCATCACTCCACTGGTTC	1682
1486	Db	GATTAGAAAATGGGTGACCCGAACACTTCTCTCATACACACCATCACTCCACTGGTTC	1545
1683	Qy	GGAGACCAGATCATGTCTCCGAGAAAGCCCTTATTTCTGTGAGCCAGAAAGTCAATCCCA	1742
1546	Db	GGAGACCAGATCATGTCTCCGAGAAAGCCCTTATTTCTGTGAGCCAGAAAGTCAATCCCA	1605
1743	Qy	AGCAGAGGCTTGCTCTGAGTCTAGAAATAGAGTAAGAGGAGGCTAGACTCAAGCTGTCT	1802
1606	Db	AGCAGAGGCTTGCTCTGAGTCTAGAAATAGAGTAAGAGGAGGCTAGACTCAAGCTGTCT	1665
1803	Qy	GGAGAGTGTGAAAACAAAGTGTGTGAAGAGTGTGTAACTGTGTGACTGTGAGCTTGAATGCCA	1862
1666	Db	GGAGAGTGTGAAAACAAAGTGTGTGAAGAGTGTGTAACTGTGTGACTGTGAGCTTGAATGCCA	1725
1863	Qy	AGTTGAAAATCTTCAATTTGGATCTGTGCTGCCCTTGTGGTAACACAGAGACACTTAGTAA	1922
1726	Db	AGTTGAAAATCTTCAATTTGGATCTGTGCTGCCCTTGTGGTAACACAGAGACACTTAGTAA	1785
1923	Qy	GGACTCTCTAGGCTCTACCAAAATCAAGCAAAAATTTGAAGAGGCTGGTACCAAGTATCTCAGA	1982
1786	Db	GGACTCTCTAGGCTCTACCAAAATCAAGCAAAAATTTGAAGAGGCTGGTACCAAGTATCTCAGA	1845
1983	Qy	GCCTCCGCTCTCTATCAGTCCGATGTGTTTCCAGAAAAGCTGTGGAAACGCTACTCTTCTTTT	2042
1846	Db	GCCTCCGCTCTCTATCAGTCCGATGTGTTTCCAGAAAAGCTGTGGAAACGCTACTCTTCTTTT	1905
2043	Qy	GAGACCTTGTGGAGAGGGTCTGAAAATGGTATGGCAAGAGAAATAGTTTCCCACAGAGATAA	2102
1906	Db	GAGACCTTGTGGAGAGGGTCTGAAAATGGTATGGCAAGAGAAATAGTTTCCCACAGAGATAA	1965
2103	Qy	AAAATGGTGTGTGGCCATGGCAGTCAAAACGGAAGGCTTGAGAATCCATCTCTCCAGAGTCC	2162
1966	Db	AAAATGGTGTGTGGCCATGGCAGTCAAAACGGAAGGCTTGAGAATCCATCTCTCCAGAGTCC	2025
2163	Qy	GTCAATCCAGACACCCAAATTTCCAGAGACAGAGCGGAAAGACATTTGCCAAGCCCGTTCAC	2222
2026	Db	GTCAATCCAGACACCCAAATTTCCAGAGACAGAGCGGAAAGAAATTTGCCAAGCCCGTTCAC	2085

```
QY 2223 CATCACGCCAGCTCCATGAGGAAATCTGCACATACCTCCATAGAAAGTCCGAGGGA 2282
Db 2086 CATCACGCCAGCTCCATGAGGAAATCTGCACATACCTCCATAGAAAGTCCGAGGGA 2145
QY 2283 CTTCTGTGGTCTGGAACACCTCAACAGAAATATAGATTTCTAATCTGAGTGATTTACTGAGC 2342
Db 2146 CTTCTGTGGTCTGGAACACCTCAACAGAAATATAGATTTCTAATCTGAGTGATTTACTGAGC 2205
QY 2343 TTTGGTCCACTAAACCAAGCTGAGCTTTGGTCCACTAAACCAAGATGAAAAATACAGAG 2402
Db 2206 TTTGGTCCACTAAACCAAGCTGAGCTTTGGTCCACTAAACCAAGATGAAAAATACAGAG 2265
QY 2403 TGAATCTATTAATCTGCTTTTAAAGAAAGCTGCTTTTCAATTTTACACAAAATCTTTTC 2462
Db 2266 TGAATCTATTAATCTGCTTTTAAAGAAAGCTGCTTTTCAATTTTACACAAAATCTTTTC 2325
QY 2463 AACGCTGAAATGATCACTTAATCTGCTTCTACTACCATTAATGATATGAGCTTCCCGAGGA 2522
Db 2326 AACGCTGAAATGATCACTTAATCTGCTTCTACTACCATTAATGATATGAGCTTCCCGAGGA 2385
QY 2523 TGAATGCTGTGTTAAATTTTCAATGATTAATTTGTCACCTCTAGCATTTTGAATGAATAG 2582
Db 2386 TGAATGCTGTGTTAAATTTTCAATGATTAATTTGTCACCTCTAGCATTTTGAATGAATAG 2445
QY 2583 TCTTTCACTTTTAAATTTTATTTCACTTCTCTATAATGATGATCCAGTTTCATGGAGGCA 2642
Db 2446 TCTTTCACTTTTAAATTTTATTTCACTTCTCTATAATGATGATCCAGTTTCATGGAGGCA 2505
QY 2643 AAAAAACAAGTTCTTGTATCTCTGAAACTTTTCTATGCTCAGTGGAAAGTATCTGCCAGCC 2702
Db 2506 AAAAAACAAGTTCTTGTATCTCTGAAACTTTTCTATGCTCAGTGGAAAGTATCTGCCAGCC 2565
QY 2703 ACAGCATGAGGCTGTGAAAGGCTGACTGAGAAATCTCTCTCTGCTGAAAGCCCTGGTCTGT 2762
Db 2566 ACAGCATGAGGCTGTGAAAGGCTGACTGAGAAATCTCTCTCTGCTGAAAGCCCTGGTCTGT 2625
QY 2763 TCTGCTCTCAACATGATATAATTTTATTTGAAATACATAATCTTTTCACTATG 2814
Db 2626 TCTGCTCTCAACATGATATAATTTTATTTGAAATACATAATCTTTTCACTATG 2677
```

RESULT 14

US-10-221-625-131

; Sequence 131, Application US/10221625

; Publication No. US20040033942A1

; GENERAL INFORMATION:

; APPLICANT: INCYTE GENOMICS, INC.

; APPLICANT: HILLMAN, Jennifer L.

; APPLICANT: BAUGHN, Mariah R.

; APPLICANT: YUE, Henry

; APPLICANT: LAL, Preeti

; APPLICANT: LU, Dyung Aina M.

; APPLICANT: PATTERSON, Chandra

; APPLICANT: AZIMZAI, Valda

; APPLICANT: BANDMAN, Olga

; APPLICANT: TANG, Y. Tom

; APPLICANT: MATHUR, Preete

; APPLICANT: SHAH, Purvi

; APPLICANT: AU-YOUNG, Janice

; APPLICANT: REDDY, Roopa

; TITLE OF INVENTION: TRANSCRIPTION FACTORS

; FILE REFERENCE: PF-0761 PCT

; CURRENT APPLICATION NUMBER: US/10/221,625

; CURRENT FILING DATE: 2001-03-13

; NUMBER OF SEQ ID NOS: 214

; SOFTWARE: PERL Program

; SEQ ID NO 131

; LENGTH: 1871

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; OTHER INFORMATION: Incyte ID No. US20040033942A1 1377380CB1

US-10-221-625-131

```
Query Match 53.7%; Score 1520; DB 8; Length 1871;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1284 CTGGCGCTTGAATAGAGGCTTAGAGGAGAAACACAGGAGGTGATAAACTTTCCACCGTGGG 1343
Db 1 CTGGCGCTTGAATAGAGGCTTAGAGGAGAAACACAGGAGGTGATAAACTTTCCACCGTGGG 60
QY 1344 TTGGGCTCTCAGAGAAAAAGAGTCAAGACTGCGCTTAGTAAACAGTAAACAGTAGGCCA 1403
Db 61 TTGGGCTCTCAGAGAAAAAGAGTCAAGACTGCGCTTAGTAAACAGTAAACAGTAGGCCA 120
QY 1404 GAGTACTCTCGCAAGCCCGGAGGTAAGTGCATCCATCCATTTCTTCCCGTCATC 1463
Db 121 GAGTACTCTCGCAAGCCCGGAGGTAAGTGCATCCATCCATTTCTTCCCGTCATC 180
QY 1464 CGCAGCTTGTGCCCAAGCTGTGCTGGAGACCTCCCTCTTCTTCAAAATCTCTACGTT 1523
Db 181 CGCAGCTTGTGCCCAAGCTGTGCTGGAGACCTCCCTCTTCTTCAAAATCTCTACGTT 240
QY 1524 CTCTATTAAAACTCTCTGCAAGCCCGGTCTCCCATCAACAGAGAGGCTCTGTCTC 1583
Db 241 CTCTATTAAAACTCTCTGCAAGCCCGGTCTCCCATCAACAGAGAGGCTCTGTCTC 300
QY 1584 CTCGCTCTCTCCAGCCACCTTTCATCTTCAAGATGTCGATTAGAACTGGGTGACCCC 1642
Db 301 CTCGCTCTCTCCAGCCACCTTTCATCTTCAAGATGTCGATTAGAACTGGGTGACCCC 360
QY 1643 GAAACACTTCTCATCACCACCATCTCACCCTGCTTCCGAGACCAAGATCATCTCTC 1702
Db 361 GAAACACTTCTCATCACCACCATCTCACCCTGCTTCCGAGACCAAGATCATCTCTC 420
QY 1703 CGAGAAAAAGCCCTTATCTCTGTGAGCCAGAGTCATCCCAAGCAGAGGCTTGTCTGAGT 1762
Db 421 CGAGAAAAAGCCCTTATCTCTGTGAGCCAGAGTCATCCCAAGCAGAGGCTTGTCTGAGT 480
QY 1763 CTAGAAATAGAGTAAGAGGAGGCTAGACTCAAGCTGTCTGAGAGAGTGGAACAAAGT 1822
Db 481 CTAGAAATAGAGTAAGAGGAGGCTAGACTCAAGCTGTCTGAGAGAGTGGAACAAAGT 540
QY 1823 GTGTGAAGAGTTGTAACCTGTGTGACTGTGCTTGTAGTGCCCAAGTGTGAAATCTTCA 1882
Db 541 GTGTGAAGAGTTGTAACCTGTGTGACTGTGCTTGTAGTGCCCAAGTGTGAAATCTTCA 600
QY 1883 ATCTGTGCTGCCCTTGTGTTAAACAGGAGACCTTTAGTAAGGACTCTCTAGGTCCTACCA 1942
Db 601 ATCTGTGCTGCCCTTGTGTTAAACAGGAGACCTTTAGTAAGGACTCTCTAGGTCCTACCA 660
QY 1943 AATCAAGCAAAATTAAGAGAGCTGGTACAGTATCTCAGAGGCTCGCTCTCTATCAGTC 2002
Db 661 AATCAAGCAAAATTAAGAGAGCTGGTACAGTATCTCAGAGGCTCGCTCTCTATCAGTC 720
QY 2003 CCTATGCTTTCAGAAAGCTGTGGAACGCTACCTCTCTTTTGAGACCTTGTGAGAGAGGTT 2062
Db 721 CGTATGCTTTCAGAAAGCTGTGGAACGCTACCTCTCTTTTGAGACCTTGTGAGAGAGGTT 780
QY 2063 CTGAAATGGTAGGCAAGAGAAATAGTTTCCCAAGAGAAATAAAACTGGTTTGTGGCCATGG 2122
Db 781 CTGAAATGGTAGGCAAGAGAAATAGTTTCCCAAGAGAAATAAAACTGGTTTGTGGCCATGG 840
QY 2123 CAGCCAAACGGAAGGCTGAGAAATCCATCTCCAGAAAGCTCGCTCATCCAGACACCAAT 2182
Db 841 CAGCCAAACGGAAGGCTGAGAAATCCATCTCCAGAAAGCTCGCTCATCCAGACACCAAT 900
QY 2183 CCAGGAGACAGAGCGGAAAGACATTGCCAAGCCCGCTCACCATCAGCCAGCTCCATCA 2242
Db 901 CCAGGAGACAGAGCGGAAAGACATTGCCAAGCCCGCTCACCATCAGCCAGCTCCATCA 960
QY 2243 GGAAATCTGCACATCTTCCATAGAAAGTCCAGGAGGACTTCTGTGCTCTGTAACACT 2302
```

```
Db 961 GGAAATCTGCACATATCTCCATAGAAAGTCCAGGAGGACTTCTGTGGTCTGAACACT 1020
Qy 2303 CAACAGAAATTATAGATTCTAAATCTGAGTGAAGTACCTGAGCTTTGGTCCACTAAAAAAGC 2362
Db 1021 CAACAGAAATTATAGATTCTAAATCTGAGTGAAGTACCTGAGCTTTGGTCCACTAAAAAAGC 1080
Qy 2363 TGAAGCTTTGGTCCACTTAAACAGAGTGAAGAAATACAAAGAGTCACTATAAATCTGTGGTCT 2422
Db 1081 TGAAGCTTTGGTCCACTTAAACAGAGTGAAGAAATACAAAGAGTCACTATAAATCTGTGGTCT 1140
Qy 2423 TTAAGAAAGCTGCTTTTCAATTTTATAGACAAAATCTTTTCAACGCTGAAATGTACTTAAT 2482
Db 1141 TTAAGAAAGCTGCTTTTCAATTTTATAGACAAAATCTTTTCAACGCTGAAATGTACTTAAT 1200
Qy 2483 CTGGTTCTACTACATAAATATATAGCATTTTGAATGAATAGTCTTCACTTTTAAATTTAT 2542
Db 1201 CTGGTTCTACTACATAAATATATAGCATTTTGAATGAATAGTCTTCACTTTTAAATTTAT 1260
Qy 2543 CATAAAGTAAATTTGTCACTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTTAT 2602
Db 1261 CATAAAGTAAATTTGTCACTAGCATTTTGAATGAATAGTCTTCACTTTTAAATTTAT 1320
Qy 2603 CATCTTCTCTATAAATATGACATCCAGTTCATGGAGGCAAAAAACAAGTTTCTTGTAT 2662
Db 1321 CATCTTCTCTATAAATATGACATCCAGTTCATGGAGGCAAAAAACAAGTTTCTTGTAT 1380
Qy 2663 CTGAAACTTTTCTATGCTCAGTGGAAAGTATCTGCCAGCCACAGCATGAGGCTGTGAAG 2722
Db 1381 CCTGAAACTTTTCTATGCTCAGTGGAAAGTATCTGCCAGCCACAGCATGAGGCTGTGAAG 1440
Qy 2723 GCTGACTGAGAAATCTCTGCTGAAGACCCCTGGTTCTGCTTCTGCTCCCAACATGTATAA 2782
Db 1441 GCTGACTGAGAAATCTCTGCTGAAGACCCCTGGTTCTGCTTCTGCTCCCAACATGTATAA 1500
Qy 2783 TTTTATTGAAATACATAATCTTTTCACTATG 2814
Db 1501 TTTTATTGAAATACATAATCTTTTCACTATG 1532

RESULT 15
US-10-330-773-670
; Sequence 670, Application US/10330773
; Publication No. US20060040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; TITLE OF INVENTION: Novel
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330, 773
; CURRENT Filing DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 670
; LENGTH: 2426
; TYPE: DNA
; ORGANISM: Mus musculus
; US-10-330-773-670

Query Match 46.0%; Score 1302.2; DB 11; Length 2426;
Best Local Similarity 87.7%; Pred. No. 0;
Matches 1434; Conservative 0; Mismatches 198; Indels 3; Gaps 1;

Qy 588 ATTCTGTACGGGTGGAGAGATGGCAACATTATGGTCTGGGATACAGGTGCAACAAAA 647
Db 629 ATGCTTCAAGGGGGGAGACGGCAACATTATGATCTGGGACACAGAGGTGTAAACAAAA 688
Qy 648 AGATGGGTTTTATAGCGCAAGTGAATCAATACGTGGGCTCAATACCTCAGACAAGCA 707
Db 689 AGATGGATTTTATAGCAAGTGAATCAATACGTGGGCTCAATACCTCAGACAAGCA 748
Qy 708 AACCCCTTCAAAACCCCAAGAGAAACAGAAATCAAAAGGACTTGTCTCTGTGGATT 767
Db 749 AACCCCTTCAAGCCCAAGAGAAACAGAAATCAAAAGGACTTGTCTCTGTGGATT 808
```

```
Qy 768 CCAGCAAGTGTCTACTGTGGTCTCTTTCAAGACGAGNAATACCTTACTCTCAGCAGGAGC 827
Db 809 CCAGCAGAGTGTCTACTGTGGTCTCTTTCAAGGATGAGAATACATTAGTCTCAGCAGGAGC 868
Qy 828 TGTGGATGGGATAATCAAAAGTATCGGAATTTACGTAAGAAATATATCTGCTTATCGACAAGA 887
Db 869 CGTGGATGGAATATCAAAAGTATCGGAATTTGCGCAAGNAATTACACTGCTTATCGACAAGA 928
Qy 888 ACCCATAGCATCCAAGTCTTTCTGTACCCAGGTAGCAGCACTCGAANAATTTGGATATTC 947
Db 929 ACCCATAGCATCCAAGTCTTTCTGTACCCAGGTAGCAGCACTCGAANAATTTAGGATATTC 988
Qy 948 AAGTCTGATTTTGGATTTCCACATGGCTCTACTTTATTTGCTAAATTTGCACAGACGATAACAT 1007
Db 989 GAGTTGGTTTTAGACTCTACTGGCTTACTTTATTTTGTAACTGCACAGATGACAACAT 1048
Qy 1008 CTACATGTTTAAATATGACTGGGTTGAAGACTTCTCCAGTGGCTATTTTCAATGGACACCA 1067
Db 1049 CTATATGTTCAATATGACTGGCTTAAAGACTTCTCCGGTGGCTGTCTTCAATGGACACCA 1108
Qy 1068 GAACTCTACCTTTTATGTAATAATCCAGCTTATGTCAGATGACAGATTTTATGTCAGTGG 1127
Db 1109 GAACTCTACCTTTTATGTAATAATCAAGTCTTAGTCCAGATGACCAGTTTTTAAATCAGTGG 1168
Qy 1128 CTCAGTGATGAAGCTGCTACATATGGAAGGTCTCCACACCCCTGGCAACCTCTCTACTGT 1187
Db 1169 TTCAGTGATGAAGCTGCTACATTTTGAAGGTTTCCATGCATGGGCATCTCTCTACTGT 1228
Qy 1188 GCTCTCGGTCATTTCTCAAGAGGTCACTCTGTGTGTGGTGTCCATCTGACTTTCACAAA 1247
Db 1229 GCTCTCGGTCATTTCTCAAGAGGTCACTCTGTGTGTGGTGTCCATCAGACTTTCACCAA 1288
Qy 1248 GATTGCTACCTGTTCTGATGACAAATACACTAAAAATCTGGCGCTTGAATAGAGGCTTAGA 1307
Db 1289 GATTGCAACCTGCTCTGATGATAATACACTGAAATCTGGCGCTTGAATAGAGGCTTAGA 1348
Qy 1308 GGAGAAACAGAGGTGATAAACTTTCCACGGTGGGTTGGGCTCTCAGAGAAAGAAAGA 1367
Db 1349 GGAGAAACCC--AGGTGATAAACATTTCCATAGTGGGTTGGACCTCTCAGAGAAAGAAAGA 1405
Qy 1368 GTCAGAGCTTGGCTTAGTAACAGATAACGAGTAGCCAGAGTACTCTCTGCAAGCCCCAG 1427
Db 1406 AGTGAAGCTTGGCCAGTAAACGGTACCAAGTAGCCAGAGTACTCTCTGCAAGCTCCAG 1465
Qy 1428 GGTAAAGTGCAATCCATCCAATTTTCCCGTCAATCCGAGCTTGTGCCCCAGCTGTGC 1487
Db 1466 AGCCAAGAGCAGTCCATCCATCTCTCTCTCTCGTCAGCAGCTTGTACTCCGAGCTGTGC 1525
Qy 1488 TGGAGACCTCCCTCTTCTCAAAATACCTCCTAGCTTCTCTATTAATAAACCCTCTCTGCCAA 1547
Db 1526 AGGAGACCTCCCTCTTCTCTCAAGTACCCCAATCTCAGTCAANAACACTCTCTGCCAC 1585
Qy 1548 GGCCCGGCTCTCCCATCAACAGAAAGGCTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1607
Db 1586 GACCCGTTCTTCAGTCAGCAGAGAGGCTCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1645
Qy 1608 ATCTTTCAAGATGTTCGATTAGAAATCTGGGTGACCCGAAACACCTTCTCTCTCTCTCTCTCT 1667
Db 1646 ATCTTTCAAGATGTTCGATTAGAAATCTGGGTGACCCGAAACACCTTCTCTCTCTCTCTCTCT 1705
Qy 1668 CACTCCACTGCTCGGAGACCAAGATCATGTCTCCGAGAAAGCCCTTATTTCTCTGTAG 1727
Db 1706 CACTCCACTGCTCTGAGACAAAGATCTCATCTCTCAAGAAAGCTCTTATTTCTCTGTAG 1765
Qy 1728 CCAGAAAGTCATCCCAAGCAGAGGCTTGTCTGTAGTCTAGAAATAGAGTAAGAGAGGCT 1787
Db 1766 CCAGAAAGTCATCAGAGGAGAGTGTCTGTCTGAAATCTAGAAATAGAGTAAGAGGCTCT 1825
Qy 1788 AGACTCAAGCTGTCTGGAGAGTGTCAACAAAGTGTGTGAAGAGTTGTAACTGTGTGAC 1847
Db 1826 TGACTCAAGCTGTCTGGAGAGTGTCAACAAAGTGTGTGAAGAGTTGTCAACTGTGTGTCAC 1885
```


Qy	1848	TGAGCTTGATGGCAAGTTCGAAATCTTCATTTGGATCTGTGCTGCTTGGTAAACCA	1907
Db	1886	TGAGCTTGACGGCCAGCGAGAGTCTTCGTTTGGATCTGTGCTTCCGGCACCCA	1945
Qy	1908	GGAAGACCTTAGTAAGACTCTCTAGTCTACCAAAATCAAGCAAAATTTGAAGGAGCTGG	1967
Db	1946	GGAAGTCTTTAGCCAGACTCCGAGGGTCTTACCAAAATCAAGCAAGACTGAAGGTGCTGG	2005
Qy	1968	TACCAGTATCTCAGAGCTCCGTCTCCTATCAGTCCGTATGCTTCAGAAAGCTGTGGAAC	2027
Db	2006	CACAAGCATCTCAGAACCTCTCTCTGTGCTTATGCTTCTGAAGGCTGTGACC	2065
Qy	2028	GCTACCTCTTCTTTGAGACCTTTGTGAGAAGGCTCTGAAATGTTAGGCAAGAGAATAG	2087
Db	2066	ACTGCCCTCTTCTTTGAGACCTTTGTGAGAAGGATCTGAGATGGTGGGCAAGAGAATAG	2125
Qy	2088	TTCCCCAGAGAATAAAACTGGTTGTGGCCATGGCAGCCAAACGGAGGCTGAGATCC	2147
Db	2126	CTCTCCAGAGAATAAGAACTGGTTGTGGCCATAGCAGCCAAACGCAAGGCAGAAAAATTC	2185
Qy	2148	ATCTCCACGAAGTCCGTCTATCCAGACACCCCAATTCAGGAGACAGAGCGGAAAGACATT	2207
Db	2186	ATCCCCAAGAGTCCATCTCAGACACCCAGTTCAGGAGACAAAGTGGGAAGACGTC	2245
Qy	2208	GCCAAGCCCGGTAC	2222
Db	2246	ACCAGGCCCGTCTAC	2260

Search completed: November 9, 2006, 07:39:38
Job time : 4573 secs

This Page Blank (uspto)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 9, 2006, 05:58:58 ; Search time 518 Seconds
(without alignments)
10226.079 Million cell updates/sec

Title: US-10-726-160-1

Perfect score: 2831

Sequence: 1 ggcacagcgggagttggag.....atgaaaaaaaaaaaaaa 2831

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters: 2807332

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /EMC Celerra_SIDS3/ptodata/2/ina/1 COMB.seq.*
- 2: /EMC Celerra_SIDS3/ptodata/2/ina/5 COMB.seq.*
- 3: /EMC Celerra_SIDS3/ptodata/2/ina/6A COMB.seq.*
- 4: /EMC Celerra_SIDS3/ptodata/2/ina/6B COMB.seq.*
- 5: /EMC Celerra_SIDS3/ptodata/2/ina/7 COMB.seq.*
- 6: /EMC Celerra_SIDS3/ptodata/2/ina/H COMB.seq.*
- 7: /EMC Celerra_SIDS3/ptodata/2/ina/PCTUS COMB.seq.*
- 8: /EMC Celerra_SIDS3/ptodata/2/ina/PP COMB.seq.*
- 9: /EMC Celerra_SIDS3/ptodata/2/ina/RE COMB.seq.*
- 10: /EMC Celerra_SIDS3/ptodata/2/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	500.4	17.7	815	4	US-09-297-648-3050 Sequence 3050, Ap
2	421.4	14.9	477	4	US-09-297-648-5115 Sequence 5115, Ap
3	300	10.6	300	4	US-09-297-648-99 Sequence 99, Appl
4	300	10.6	300	4	US-09-297-648-1050 Sequence 1050, Ap
5	290.4	10.3	346	4	US-09-880-107-3634 Sequence 3634, Ap
6	276.8	9.8	290	3	US-09-513-999C-21911 Sequence 21911, A
7	156	5.5	413	3	US-09-513-999C-21244 Sequence 21244, A
8	53	1.9	7218	2	US-08-232-463-14 Sequence 14, Appl
9	48.2	1.7	1548	3	US-09-614-221A-136 Sequence 136, App
10	44.2	1.6	7218	2	US-08-232-463-14 Sequence 14, Appl
11	44	1.6	832	3	US-09-621-976-2813 Sequence 2813, Ap
12	42.8	1.5	1141	3	US-09-806-708B-22 Sequence 22, Appl
13	40.8	1.4	2025	3	US-09-252-991A-1536 Sequence 1536, Ap
14	40.6	1.4	77851	3	US-09-949-016-12508 Sequence 12508, A
15	40.6	1.4	77857	3	US-09-949-016-13211 Sequence 13211, A
16	40.6	1.4	77857	3	US-09-949-016-13212 Sequence 13212, A
17	40.6	1.4	77940	3	US-09-949-016-12509 Sequence 12509, A
18	40.4	1.4	84525	3	US-09-949-016-16578 Sequence 16578, A
19	40.2	1.4	1141	3	US-09-806-708B-22 Sequence 22, Appl
20	40	1.4	601	3	US-09-949-016-174701 Sequence 174701, A
21	39.6	1.4	63658	3	US-09-949-016-13238 Sequence 13238, A
22	39.6	1.4	64489	3	US-09-949-016-11766 Sequence 11766, A
23	38.8	1.4	700	3	US-09-735-271-383 Sequence 383, App

Sequence 6297, Ap
Sequence 2, Appli
Sequence 2, Appli
Sequence 996, App
Sequence 3156, Ap
Sequence 18438, A
Sequence 1433, Ap
Sequence 16480, A
Sequence 211, App
Sequence 36, Appl
Sequence 36, Appl
Sequence 36, Appl
Sequence 36, Appl
Sequence 36, Appl
Sequence 36, Appl
Sequence 1960, Ap
Sequence 266, App
Sequence 4891, Ap
Sequence 897, App

24 38.8 1.4 1593 3 US-09-248-796A-6297
25 38.6 1.4 1211 3 US-09-063-743-2
26 38.6 1.4 1211 3 US-09-590-540-2
27 38 1.3 2724 4 US-10-094-749-996
28 37.8 1.3 561 3 US-09-270-767-3156
29 37.8 1.3 561 3 US-09-270-767-18438
30 37.8 1.3 1230 3 US-09-252-991A-1433
31 37.6 1.3 225127 3 US-09-949-016-16480
32 37.6 1.3 1082144 4 US-09-531-120-211
33 37.4 1.3 475 3 US-09-702-705-36
34 37.4 1.3 475 3 US-09-736-457-36
35 37.4 1.3 475 3 US-09-614-124B-36
36 37.4 1.3 475 3 US-09-671-325-36
37 37.4 1.3 475 3 US-09-589-184-36
38 37.4 1.3 475 3 US-09-658-824-36
39 37.4 1.3 475 3 US-10-017-754-36
40 37.4 1.3 475 3 US-09-651-563-36
41 37.4 1.3 475 3 US-09-519-642-36
42 37.4 1.3 744 3 US-09-248-796A-1960
43 37.4 1.3 1262 3 US-09-149-476-266
44 37.4 1.3 1431 3 US-09-248-796A-4891
45 37.4 1.3 6158 3 US-09-799-451-897

ALIGNMENTS

RESULT 1

US-09-297-648-3050
; Sequence 3050, Application US/09297648
; Patent No. 6964868
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Escobedo, Jaime A.
; APPLICANT: Innis, Michael A.
; APPLICANT: Garcia, Pablo Dominguez
; APPLICANT: Sudduth-Klinger, Julie
; APPLICANT: Reinhard, Christoph
; APPLICANT: Giese, Klaus
; APPLICANT: Randazzo, Filippo
; APPLICANT: Kennedy, Giulia C.
; APPLICANT: Pot, David
; APPLICANT: Kassan, Altaf
; APPLICANT: Lamson, George
; APPLICANT: Drmanac, Radoje
; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Garcia, Veronica
; APPLICANT: Jones, William Lee
; APPLICANT: Stache-Crain, Birjit
; TITLE OF INVENTION: No. 6964868 Human Genes and Gene Expression
; FILE REFERENCE: 2300-1481
; CURRENT APPLICATION NUMBER: US/09/297, 648
; CURRENT FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/072,910
; PRIOR FILING DATE: 1998-01-28
; PRIOR APPLICATION NUMBER: 60/075,954
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/080,666
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/080,515
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/080,114
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/105,234
; PRIOR FILING DATE: 1998-10-21
; NUMBER OF SEQ ID NOS: 5252
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3050
; LENGTH: 815
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(815)
; OTHER INFORMATION: n = A,T,C or G
US-09-297-648-3050

Query Match
Best Local Similarity 17.7%; Score 500.4; DB 4; Length 815;
Matches 559; Conservative 0; Mismatches 13; Indels 5; Gaps 5;
QY 1780 AGGAGCTAGACTCAAGCTGTCTGGAGAGTGTGAACAAAGTGTGAAGAGTTGTAAAC 1839
Db 65 ACGAGCTAGACTCAAGCTGTCTGGAGAGTGTGAACAAAGTGTGAAGAGTTGTAAAC 124
QY 1840 TGTGTGACTGAGCTTGATGCCCAAGTTGAAATCTTCATTTGGATCTGTGCTGCCTTGGT 1899
Db 125 TGTGTGACTGAGCTTGATGCCCAAGTTGAAATCTTCATTTGGATCTGTGCTGCCTTGGT 184
QY 1900 GGTAAACAGAGACCTTAGTAAGGACTCTTAGGTCTCTAGGTCTTACCAATCAAGCAAAATTGAA 1959
Db 185 GGTAAACAGAGACCTTAGTAAGGACTCTTAGGTCTTACCAATCAAGCAAAATTGAA 244
QY 1960 GGAGCTGGTACCAGTATCTCAGAGCTCCGTCTCCTATCAGTCCGTATGCTTCAGAAAGC 2019
Db 245 GGAGCTGGTACCAGTATCTCAGAGCTCCGTCTCCTATCAGTCCGTATGCTTCAGAAAGC 304
QY 2020 TGTGGACCTACCTCTTCTTTGAGACCTTGTGAGAGAGGTCTGAAATGTAGGCATA 2079
Db 305 TGTGGACCTACCTCTTCTTTGAGACCTTGTGAGAGAGGTCTGAAATGTAGGCATA 364
QY 2080 GAGATAGTTCCCAAGAGATAAAACTGTTGTTGGCCATGCGACCAACCGAAGGCT 2139
Db 365 GAGATAGTTCCCAAGAGATAAAACTGTTGTTGGCCATGCGACCAACCGAAGGCT 424
QY 2140 GAGATCCATCTCCAGAGTCCGTATCCCAAGACACCCAAATCCAGGAGACAGAGCGGA 2199
Db 425 GAGATCCATCTCCAGAGTCCGTATCCCAAGACACCCAAATCCAGGAGACAGAGCGGA 484
QY 2200 AAGACATCCCAAGCCCGTCCACCATCAGCCCGAGTCCATGAGGAAATCTGCACATAC 2259
Db 485 AAGACATTCNCAAGCCCGG-CACCATCAGCCCGAGTCCATGAGGAAATCTGCACATAC 543
QY 2260 TTCCATAGAAAGTCCCAAGAGAC-TTCGTGTCTCTGAACACTCAA-CAGAAATTATA-G 2316
Db 544 TTCCATAGAAAGTCCCAAGAGAC-TTCGTGTCTCTGAACACTCAA-CAGAAATTATA-G 603
QY 2317 ATTCTAATCTGAG-TGAGTTACTGAGCTTTTGGTCCAC 2352
Db 604 ATTCTAATCTGAGTTGAGTTACTGAGCTTTTGGTCCC 640

RESULT 2

US-09-297-648-5115
; Sequence 5115, Application US/09297648
; Patent No. 6964868
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Escobedo, Jaime
; APPLICANT: Innis, Michael A.
; APPLICANT: Garcia, Pablo Dominiguez
; APPLICANT: Sudduth-Klinger, Julie
; APPLICANT: Reinhard, Christoph
; APPLICANT: Giese, Klaus
; APPLICANT: Randazzo, Filippo
; APPLICANT: Kennedy, Giulia C.
; APPLICANT: Pot, David
; APPLICANT: Kassan, Altaf
; APPLICANT: Lamson, George
; APPLICANT: Drmanac, Radoje

; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Garcia, Veronica
; APPLICANT: Jones, William Lee
; APPLICANT: Stache-Crain, Birjit
; TITLE OF INVENTION: No. 6964868el Human Genes and Gene Expression
; FILE REFERENCE: 2300-1481
; CURRENT APPLICATION NUMBER: US/09/297,648
; PRIOR FILING DATE: 2000-03-10
; PRIOR APPLICATION NUMBER: 60/072,910
; PRIOR FILING DATE: 1998-01-28
; PRIOR APPLICATION NUMBER: 60/075,954
; PRIOR FILING DATE: 1998-02-24
; PRIOR APPLICATION NUMBER: 60/080,666
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/080,515
; PRIOR FILING DATE: 1998-04-03
; PRIOR APPLICATION NUMBER: 60/080,114
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/105,234
; PRIOR FILING DATE: 1998-10-21
; NUMBER OF SEQ ID NOS: 5252
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5115
; LENGTH: 477
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-297-648-5115

Query Match 14.9%; Score 421.4; DB 4; Length 477;
Best Local Similarity 99.5%; Pred. No. 4.7e-116;
Matches 433; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1785 GCTAGACTCAAGCTGTCTGGAGAGTGTGAACAAAGTGTGTGAAGAGTTGTAACTGTGT 1844
Db 1 GCTAGACTCAAGCTGTCTGGAGAGTGTGAACAAAGTGTGTGAAGAGTTGTAACTGTGT 60
QY 1845 GACTGAGCTTGTGTCGCAAGTGTGAACATCTTCATTTGGATCTGTGCTGCTTGGTAA 1904
Db 61 GACTGAGCTTGTGTCGCAAGTGTGAACATCTTCATTTGGATCTGTGCTGCTTGGTAA 120
QY 1905 CCAGGAAGACCTTAGTAAGGACTCTCTAGGTCTTACCAAAATCAAGCAAAATTTGAAGGAGC 1964
Db 121 CCAGGAAGACCTTAGTAAGGACTCTCTAGGTCTTACCAAAATCAAGCAAAATTTGAAGGAGC 180
QY 1965 TGGTACCAGTATCTCAGAGCTCCGTCTCCTATCAGTCCGTATGCTTCAGAAAGCTGTGG 2024
Db 181 TGGTACCAGTATCTCAGAGCTCCGTCTCCTATCAGTCCGTATGCTTCAGAAAGCTGTGG 240
QY 2025 AACGCTACCTCTTCCCTTTGAGACCTTTGGAGAGAGGTCTGAAATGGTAGGCAAGAGAA 2084
Db 241 AACGCTACCTCTTCCCTTTGAGACCTTTGGAGAGAGGTCTGAAATGGTAGGCAAGAGAA 300
QY 2085 TAGTTCCCAAGAGATAAAACTGCTTGTGGCCATGGCAGCAACCGAAGGCTGAGAA 2144
Db 301 TAGTTCCCAAGAGATAAAACTGCTTGTGGCCATGGCAGCAACCGAAGGCTGAGAA 359
QY 2145 TCCATCTCAGAAAGTCCGTATCCAGACACCCCAATTTCCAGGAGACAGAGCGGAAAGAC 2204
Db 360 TCCATCTCAGAAAGTCCGTATCCAGACACCCCAATTTCCAGGAGACAGAGCGGAAAGAC 419
QY 2205 ATTGCCAAGCCCGGT 2219
Db 420 ATTGCCAAGCCCGCT 434
RESULT 3
US-09-297-648-99

; Sequence 99, Application US/09297648					
; Patent No. 6964868					
; GENERAL INFORMATION:					
; APPLICANT: Williams, Lewis T.					
; APPLICANT: Escobedo, Jaime					
; APPLICANT: Innis, Michael A.					
; APPLICANT: Garcia, Pablo Dominiguez					
; APPLICANT: Reinhard, Christoph					
; APPLICANT: Giese, Klaus					
; APPLICANT: Randazzo, Filippo					
; APPLICANT: Kennedy, Giulia C.					
; APPLICANT: Pot, David					
; APPLICANT: Kassan, Altaf					
; APPLICANT: Lamson, George					
; APPLICANT: Drmanac, Radoje					
; APPLICANT: Crkvenjakov, Radomir					
; APPLICANT: Dickson, Mark					
; APPLICANT: Labat, Ivan					
; APPLICANT: Leshkowitz, Dena					
; APPLICANT: Kita, David					
; APPLICANT: Jones, William Lee					
; APPLICANT: Stache-Crain, Birjit					
; TITLE OF INVENTION: No. 6964868el Human Genes and Gene Expression					
; FILE REFERENCE: 2300-1481					
; CURRENT APPLICATION NUMBER: US/09/297,648					
; PRIOR FILING DATE: 2000-03-10					
; PRIOR FILING DATE: 1998-01-28					
; PRIOR FILING DATE: 1998-02-24					
; PRIOR FILING DATE: 1998-04-03					
; PRIOR FILING DATE: 1998-03-31					
; PRIOR FILING DATE: 1998-10-21					
; NUMBER OF SEQ ID NOS: 5252					
; SOFTWARE: FastSeq for Windows Version 4.0					
; SEQ ID NO 99					
; LENGTH: 300					
; TYPE: DNA					
; ORGANISM: Homo sapiens					
US-09-297-648-99					
Query Match 10.6%; Score 300; DB 4; Length 300;					
Best Local Similarity 100.0%; Pred. No. 1.2e-79;					
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1785	GCTAGCACTCAAGTCTCTGGAGAGTGTAACAAAGTTGTGGAAGAGTTGTAAGTGTG	1844		
DB	1	GCTAGCACTCAAGTCTCTGGAGAGTGTAACAAAGTTGTGGAAGAGTTGTAAGTGTG	60		
QY	1845	GACTCAGCTTGATGCCAAGTTGAAAATCTTTTCATTGCGATCTGTCGCCTTCCTGATA	1904		
DB	61	GACTCAGCTTGATGCCAAGTTGAAAATCTTTTCATTGCGATCTGTCGCCTTCCTGATA	120		
QY	1905	CAGAAGACCTTAGTAGGACTCTAGTCTCCATCCAAATCAAGCAAAATTTAGAGGC	1964		
DB	121	CAGAAGACCTTAGTAGGACTCTAGTCTCCATCCAAATCAAGCAAAATTTAGAGGC	180		
QY	1965	TGGTACCAGTATCTCAGAGCCTCCGCTCTCATCAGTCCGATGCTCTCAGAAAGCTGTGG	2024		
DB	181	TGGTACCAGTATCTCAGAGCCTCCGCTCTCATCAGTCCGATGCTCTCAGAAAGCTGTGG	240		
QY	2025	AACGCTACTCTCTCTTTGAGACCTTTGGAGAGGGTCTGAAATGGTAGGCAAGAGAA	2084		
DB	241	AACGCTACTCTCTCTTTGAGACCTTTGGAGAGGGTCTGAAATGGTAGGCAAGAGAA	300		

; Sequence 99, Application US/09297648					
; Patent No. 6964868					
; GENERAL INFORMATION:					
; APPLICANT: Williams, Lewis T.					
; APPLICANT: Escobedo, Jaime					
; APPLICANT: Innis, Michael A.					
; APPLICANT: Garcia, Pablo Dominiguez					
; APPLICANT: Reinhard, Christoph					
; APPLICANT: Giese, Klaus					
; APPLICANT: Randazzo, Filippo					
; APPLICANT: Kennedy, Giulia C.					
; APPLICANT: Pot, David					
; APPLICANT: Kassan, Altaf					
; APPLICANT: Lamson, George					
; APPLICANT: Drmanac, Radoje					
; APPLICANT: Crkvenjakov, Radomir					
; APPLICANT: Dickson, Mark					
; APPLICANT: Labat, Ivan					
; APPLICANT: Leshkowitz, Dena					
; APPLICANT: Kita, David					
; APPLICANT: Garcia, Veronica					
; APPLICANT: Jones, William Lee					
; APPLICANT: Stache-Crain, Birjit					
; TITLE OF INVENTION: No. 6964868el Human Genes and Gene Expression					
; FILE REFERENCE: 2300-1481					
; CURRENT APPLICATION NUMBER: US/09/297,648					
; PRIOR FILING DATE: 2000-03-10					
; PRIOR FILING DATE: 1998-01-28					
; PRIOR FILING DATE: 1998-02-24					
; PRIOR FILING DATE: 1998-04-03					
; PRIOR FILING DATE: 1998-03-31					
; PRIOR FILING DATE: 1998-10-21					
; NUMBER OF SEQ ID NOS: 5252					
; SOFTWARE: FastSeq for Windows Version 4.0					
; SEQ ID NO 99					
; LENGTH: 300					
; TYPE: DNA					
; ORGANISM: Homo sapiens					
US-09-297-648-99					
Query Match 10.6%; Score 300; DB 4; Length 300;					
Best Local Similarity 100.0%; Pred. No. 1.2e-79;					
Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1785	GCTAGCACTCAAGTCTCTGGAGAGTGTAACAAAGTTGTGGAAGAGTTGTAAGTGTG	1844		
DB	1	GCTAGCACTCAAGTCTCTGGAGAGTGTAACAAAGTTGTGGAAGAGTTGTAAGTGTG	60		
QY	1845	GACTCAGCTTGATGCCAAGTTGAAAATCTTTTCATTGCGATCTGTCGCCTTCCTGATA	1904		
DB	61	GACTCAGCTTGATGCCAAGTTGAAAATCTTTTCATTGCGATCTGTCGCCTTCCTGATA	120		
QY	1905	CAGAAGACCTTAGTAGGACTCTAGTCTCCATCCAAATCAAGCAAAATTTAGAGGC	1964		
DB	121	CAGAAGACCTTAGTAGGACTCTAGTCTCCATCCAAATCAAGCAAAATTTAGAGGC	180		
QY	1965	TGGTACCAGTATCTCAGAGCCTCCGCTCTCATCAGTCCGATGCTCTCAGAAAGCTGTGG	2024		
DB	181	TGGTACCAGTATCTCAGAGCCTCCGCTCTCATCAGTCCGATGCTCTCAGAAAGCTGTGG	240		
QY	2025	AACGCTACTCTCTCTTTGAGACCTTTGGAGAGGGTCTGAAATGGTAGGCAAGAGAA	2084		
DB	241	AACGCTACTCTCTCTTTGAGACCTTTGGAGAGGGTCTGAAATGGTAGGCAAGAGAA	300		


```

RESULT 7
US-09-513-999C-21244
; Sequence 21244, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

```

```

? PATENT NO: 0702701
? FILE REFERENCE: 59.US2.REG
? CURRENT APPLICATION NUMBER: US/09/513,999C
? CURRENT FILING DATE: 2000-02-24
? PRIOR APPLICATION NUMBER: US 60/122,487
? PRIOR FILING DATE: 1999-02-26
? NUMBER OF SEQ ID NOS: 36681
? SOFTWARE: Patent.pm
? SEQ ID NO 21244
? LENGTH: 413
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 8
? OTHER INFORMATION: n=a, g, c or t
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 9
? OTHER INFORMATION: h=a or c or t
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 10
? OTHER INFORMATION: m=a or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 14
? OTHER INFORMATION: m=a or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 66
? OTHER INFORMATION: r=a or g
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 73
? OTHER INFORMATION: s=g or c
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 159
? OTHER INFORMATION: k=g or t
? US-09-513-999C-21244

```

Query Match	5.5%	Score 156;	DB 3;	Length 413;
Best Local Similarity	96.0%;	Pred. No. 3.2e-36;		
Matches 167;	Conservative 3;	Mismatches 167;	Indels 1;	Gaps 1;
Qy	8	GCGGAGCTGGAGGCGATAACGATTTCTGTTGTGAGAGGCGACAGTCGATTTCTGCTG	67	
Db	29	GCGGAGCTGGAGGCGATAACGATTTGTTGTGAGAGCGCAA-STGCGATTTCTGCTG	87	
Qy	68	AACCTTGAGGCGATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGC	127	
Db	88	AACCTTGAGGCGATTTCTACGACTTTTCTCTCAGCTGAGGCTTTTCTCCGACCCCTGATGC	147	
Qy	128	TCTTCAATTCGGTGCTCGCGCAGCCCGACGCTTGGGCTCTCAGAAATGGATGGT	181	
Db	148	TCTTCAATTCGGTGCTCGCGCAGCCCGACGCTTGGGCTCTCAGAAATGGATGGT	201	

RESULT 8
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:

APPLICANT: DORNER, F.
APPLICANT: SCHEIFELINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-F18
US-08-232-463-14

[illegible]

```
RESULT 9
US-09-614-221A-136
; Sequence 136, Application US/09614221A
; Patent No. 6723837
; GENERAL INFORMATION:
; APPLICANT: Karunananda, Balasulojini
; APPLICANT: Yu, Jaehyuk
; APPLICANT: Kishore, Ganesh M.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED
; WITH STRANDEDNESS: WITH STEROL SYNTHESIS AND METABOLISM
; FILE REFERENCE: 16516.075
; CURRENT APPLICATION NUMBER: US/09/614,221A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/142,981
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 626
; SEQ ID NO 136
; LENGTH: 1548
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-614-221A-136

Query Match 1.7%; Score 48.2; DB 3; Length 1548;
Best Local Similarity 55.8%; Pred. No. 0.0021;
Matches 92; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 405 GATGGCTACTGGAATGCCGCTTTGACCTGCGCTGGGTTCTCTGGTGAACCTTAAACTTGT 464
Db 1194 GACCGCTACCAAAATTAGTCAATCATGTGGCTTCAGCCCTGATGGTAGGTATATTGT 1253

QY 465 TACAGCAGCAGGTATCAACAGCCAAATTTGGGACGTAAAGCTGGTGAAGCTGATGG 524
Db 1254 CTCAGCGTCTTTGATAACTTATCAAACTTTGGGACGTAGAGATGGTAAGTTTATCTC 1313

QY 525 AACATCAAGGTCAATCAATGACGCTCAAGTCAGTTGCTTTTC 569
Db 1314 CACATTAGGGCATATAGCCAGGTATACAGGTTGGCTGGTC 1358

RESULT 10
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
```

```
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match 1.6%; Score 44.2; DB 2; Length 7218;
Best Local Similarity 4.1%; Pred. No. 0.091;
Matches 16; Conservative 213; Mismatches 166; Indels 0; Gaps 0;

QY 382 AGAAGAAGTCCTTCAAGAATGATGGCTCACTGGAATCGCGTCTTTGACCTGCGCTGG 441
Db 1432 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1373

QY 442 GTTCTCGTGAACCTTAAACTTGTACAGCAGCAGTGATCAACAGCAGCAAAATTTTGGGAC 501
Db 1372 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1313

QY 502 GTAAAAGCTGGTGAAGTGAATGGAACATGCAAGAGTCAATCAATGACGCTCAAGTCAGTT 561
Db 1312 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1253

QY 562 GCCTTTTCTAAGTTGAGAAAGCTGTATTCTGTACGGTGAAGAGATGCAACATTATG 621
Db 1252 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1193

QY 622 GTCTGGGATACAGTGCAACAAAGATGGGTTTATAGGCAAGTGAATCAAAATCAGT 681
Db 1192 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1133

QY 682 GGAGCTCAATACCTCAGACAGCAACACCCCTTCAAAACCCAGAGAGAGAAAGAGAAATCA 741
Db 1132 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1073

QY 742 AAAGGACTTGCTCTCTGTTGATTTCCAGCAAG 776
Db 1072 RRRRRATCGCAAGCTCCCTCGACCTGACGCAAG 1038

RESULT 11
US-09-621-976-2813/c
; Sequence 2813, Application US/09621576
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.E.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
US-09-621-976-2813

Query Match 1.6%; Score 44; DB 3; Length 832;
Best Local Similarity 12.3%; Pred. No. 0.026;
```


